

分析师：

徐寅

xuyinsh@xyzq.com.cn

S0190514070004

研究助理：

郑兆磊

zhengzhaolei@xyzq.com.cn

基于集成学习算法的量化选股模型研究

2019 年 06 月 21 日

报告关键点

我们引入改进后的 Adaboost 算法，以过去 12 个月的因子表现与股票月收益为训练样本，在考虑了行业市值中性化及失败者样本的情形下，构建了 E-NELS 因子。该因子 IC 为 0.136，年化 IC_IR 达到了 7.0，多空策略年化收益率 61.2%，夏普率高达 6.12，最大回撤仅为 0.5%，有效性非常强。进一步地，我们根据市场投资者的偏好，构建了换手率处于不同档位的选股模型。以高换手的周度选股模型为例，其因子 IC 达到 0.12，t 值高达 33；多空策略年化收益率达到 124.5%，夏普率高达 9.82。

相关报告

投资要点

- 四年前，我们通过 5 篇报告系统性的介绍了机器学习算法在量化选股领域的应用；四年后，我们以改进 Adaboost 算法为核心，扩展后的兴业定量因子库为基础，构建了动态选股模型。改变的是视角（由广度过渡到深度），不变的是初衷，我们构建的选股体系依然新颖有效。
- 我们引入改进后的 Adaboost 算法，以过去 12 个月的因子表现与股票月收益为训练样本，在考虑了行业市值中性化及失败者样本的情形下，构建了 E-NELS 因子。该因子 IC 为 0.136，年化 IC_IR 达到了 7.0，多空策略年化收益率 61.2%，夏普率高达 6.12，最大回撤仅为 0.5%，有效性非常强。
- 进一步地，我们根据市场投资者的偏好，构建了换手率处于不同档位的选股模型。以高换手的周度选股模型为例，其因子 IC 达到 0.12，t 值高达 33；多空策略年化收益率达到 124.5%，夏普率高达 9.82。
- 我们基于月度及周度 E-NELS 因子构建了模拟组合策略体系，包括了指数增强和主动量化在内的 11 个模型，整体表现优异。以针对于中证 500 的月度增强策略为例：超额收益率为 12.6%，信息比率为 4.27，最大回撤仅为 2.1%。

团队成员：

风险提示：本报告模型及结论全部基于对历史数据的分析，当市场环境变化时，存在模型失效风险。

目 录

1、引言	5 -
1.1、人工智能与量化投资	5 -
1.2、机器学习之集成学习	7 -
1.3、历史工作回顾	9 -
2、基于 Boosting 集成学习的量化选股模型	10 -
2.1、Adaboost 算法介绍	10 -
2.2、基于改进 Adaboost 算法的量化选股模型	13 -
2.3、ELS 因子表现	15 -
2.4、行业市值中性化	16 -
2.5、模型优势	18 -
2.6、玩转风格轮动	23 -
3、模型改进与增强	24 -
3.1、失败乃成功之母	24 -
3.2、月度效应	28 -
3.3、总结	30 -
4、高低换手风格选股模型搭建	30 -
4.1、低换手模型构建	30 -
4.2、高换手模型构建	33 -
5、股票模拟组合体系搭建	35 -
5.1、因子在沪深 300&中证 500 股票池内选股表现	35 -
5.2、基于月度 E-NELS 因子的增强策略研究	36 -
5.3、基于月度 E-NELS 因子的主动量化策略研究	40 -
5.4、基于周度 E-NELS 因子的增强模型研究	41 -
5.5、基于周度 E-NELS 因子的主动量化策略研究	42 -
5.6、模拟组合表现汇总	44 -
6、总结	45 -
7、附录	46 -

图表 1、人工智能发展的历程	- 5 -
图表 2、人工智能在量化投资领域的应用个案	- 6 -
图表 3、人工智能在投资领域的部分应用案例	- 6 -
图表 4、机器学习算法分类	- 8 -
图表 5、Bagging Vs Boosting 流程	- 9 -
图表 6、过往研究回顾	- 10 -
图表 7、股权集中度因子分位数测试效果	- 12 -
图表 8、改进 Adaboost 算法构建选股模型流程图	- 14 -
图表 9、兴业量化选股因子分类数量统计	- 14 -
图表 10、ELS 因子 IC 测试结果	- 16 -
图表 11、ELS 因子分位数组合测试结果	- 16 -
图表 12、ELS 因子多头以及多空组合表现	- 16 -
图表 13、NELS Vs.ELS 因子 IC 测试	- 17 -
图表 14、NELS Vs.ELS 因子 IC 历史走势	- 17 -
图表 15、NELS Vs.ELS 分位数组合测试结果	- 17 -
图表 16、NELS Vs.ELS 分位组合年化收益和夏普率对比	- 18 -
图表 17、NELS 因子的多空净值曲线	- 18 -
图表 18、修正后的股权集中度因子分位数测试结果	- 19 -
图表 19、因子重要性程度示意图	- 19 -
图表 20、与 EP_FY1 相关性高的因子展示	- 20 -
图表 21、弱分类器个数敏感性分析之 IC	- 21 -
图表 22、弱分类器个数敏感性分析之 IC_IR	- 21 -
图表 23、弱分类器个数敏感性分析之多空组合年化收益率	- 21 -
图表 24、弱分类器个数敏感性分析之多空组合夏普率	- 22 -
图表 25、NELS 因子的选股命中率	- 22 -
图表 26、NELS 构成因子的风格分析（一）	- 23 -
图表 27、NELS 构成因子的风格分析（二）	- 24 -
图表 28、L-NELS 因子的 IC 检验	- 25 -
图表 29、L-NELS 因子分位数组合测试结果	- 25 -
图表 30、L-NELS 因子多空净值曲线	- 25 -
图表 31、NELS & L-NELS 因子在不同月份的 IC 表现	- 26 -
图表 32、NELS 与 E-NELS 因子在不同情况下合成的 IC 检验	- 26 -
图表 33、NELS 与 E-NELS 因子在不同情况下合成的 IC 月度均值	- 26 -
图表 34、E-NELS 分位数组合测试结果	- 27 -
图表 35、E-NELS 因子多空净值曲线	- 27 -
图表 36、E-NELS 因子与原有因子对比之 IC 分析	- 27 -
图表 37、E-NELS 因子与原有因子对比之分位数测试分析	- 28 -
图表 38、NELS 因子月度 IC 表现	- 28 -
图表 39、NELS 以及 NELS_Mth 因子 IC 检验	- 29 -
图表 40、NELS 以及 NELS_Mth 因子的月度 IC 表现	- 29 -
图表 41、NELS 以及 NELS_Feb_Rep 因子 IC 对比	- 29 -
图表 42、NELS 以及 NELS_Feb_Rep 因子的多空净值曲线	- 30 -
图表 43、不同风格因子的自相关性系数均值	- 31 -
图表 44、不同阈值下的因子池总数	- 31 -
图表 45、低换手模型对阈值敏感性分析之 IC 检验	- 31 -
图表 46、低换手模型对阈值敏感性分析之自相关性以及多头换手率	- 32 -
图表 47、预测窗口延长下的低换手因子 IC 表现	- 32 -
图表 48、预测窗口延长下的低换手因子自相关性以及多头换手率	- 32 -
图表 49、周度选股因子的 IC 敏感性分析	- 33 -

图表 50、周度模型的弱分类器敏感性分析之 IC 测试.....	33 -
图表 51、周度模型的弱分类器敏感性分析之多空表现.....	34 -
图表 52、周度因子的 IC 检验	34 -
图表 53、NELS_Week Vs. E-NELS_Week 因子分位数组合测试结果.....	34 -
图表 54、E-NELS_Week 因子的多空净值曲线（对数化）	35 -
图表 55、策略层面的研究框架	35 -
图表 56、周度&月度 E-NELS 因子在沪深 300 及中证 500 股票池内 IC 测试.....	36 -
图表 57、月度 E-NELS 因子在沪深 300 及中证 500 股票池内分位数测试.....	36 -
图表 58、周度 E-NELS 因子在沪深 300 及中证 500 股票池内分位数测试.....	36 -
图表 59、基于月度 E-NELS 因子构建的增强策略框架.....	37 -
图表 60、针对于沪深 300 的增强策略表现汇总.....	37 -
图表 61、针对于沪深 300 构建的增强策略分年度表现.....	37 -
图表 62、针对于沪深 300 构建的增强策略多头净值曲线.....	38 -
图表 63、针对于沪深 300 构建的增强策略超额净值曲线.....	38 -
图表 64、针对于中证 500 的增强策略表现汇总.....	38 -
图表 65、针对于中证 500 构建的增强策略分年度表现.....	39 -
图表 66、针对于中证 500 构建的增强策略多头净值曲线.....	39 -
图表 67、针对于中证 500 构建的增强策略超额净值曲线.....	39 -
图表 68、基于月度 E-NELS 因子构建的主动量化策略.....	40 -
图表 69、月度主动量化策略表现汇总	40 -
图表 70、月度主动量化策略分年度表现	40 -
图表 71、月度主动量化策略多头净值曲线	41 -
图表 72、月度主动量化策略超额净值曲线	41 -
图表 73、周度 E-NELS 因子增强策略整体表现.....	42 -
图表 74、周度 E-NELS 因子增强策略净值曲线.....	42 -
图表 75、周度 E-NELS 因子增强策略分年度表现.....	42 -
图表 76、基于周度 E-NELS 因子构建的主动量化策略框架.....	42 -
图表 77、针对于中证 500 的周度主动量化策略表现汇总.....	43 -
图表 78、针对于中证 500 构建的周度主动量化策略分年度表现.....	43 -
图表 79、针对于中证 500 构建的周度主动量化策略多头净值曲线.....	43 -
图表 80、针对于中证 500 构建的周度主动量化策略超额净值曲线.....	44 -
图表 81、基于月度 E-NELS 因子构建的策略表现回顾.....	44 -
图表 82、基于周度 E-NELS 因子构建的策略表现回顾.....	44 -
图表 83、部分因子释义	46 -

报告正文

1、引言

1.1、人工智能与量化投资

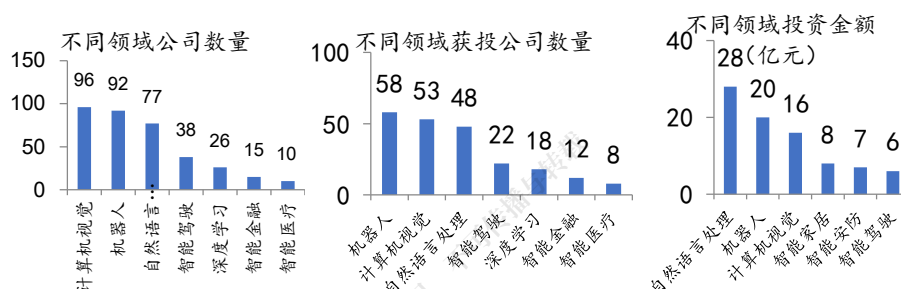
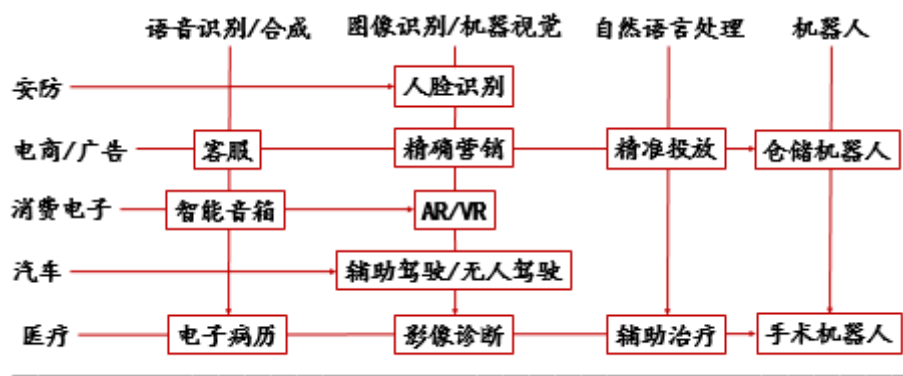
1956 年的夏天，一场在美国达特茅斯（Dartmouth）大学召开的学术会议，多年以后被认定为全球人工智能研究的起点。1968 年斯坦利库布里克作品《2001：太空漫游》里的 HAL9000，到最近《星战：原力觉醒》的 BB-8……电影让我们更真实的了解人工智能的魅力。从 Alpha Go 跟李世石围棋博弈技惊四座，到柯洁的再次挑战……现实让我们意识到人工智能就在你我身边，并不遥远。是的，人工智能就在你我身边：每天上网完成的搜索、扫地机器人、智能冰箱、车载导航、你我每天离不开的智能手机……人工智能也终将彻底的、毫无保留的改变这个世界，包括量化投资领域。

图表 1、人工智能发展的历程



资料来源：兴业证券经济与金融研究院整理

图表 2、人工智能在量化投资领域的应用个案



备注：统计时点：2017 年 12 月 31 日

资料来源：兴业证券经济与金融研究院整理

自从上世纪七十年代发端于美国，量化投资已经走过了四十多年的发展历程。近些年来，随着进入这一领域的资金和人才不断出现井喷式的增长，传统量化投资者，特别是以使用多因子框架为代表的股票量化投资管理人，其所面临的挑战和困难也愈发凸显出来：现有的因子和策略体系已经被过度挖掘和开发，推陈出新、获取超额收益的难度不断加大。更令人担忧的是，由于研究方法和所使用的数据同质性非常高，策略的拥挤程度与日俱增，一旦出现流动性风险或是事先无法预料的大规模清算，那么类似于 2007 年 8 月份发生的“量化危机”也必将重现。

股票量化投资管理的未来会走向何方？也许每个人都有自己的思考，而本文正是我们团队回答上述问题的一次尝试。在人工智能发展方兴未艾的今天，我们尝试将量化投资与人工智能的重要子领域——机器学习有机结合，我们期望两者的碰撞能为投资业带来不一样的风景。而这种看似跨界的融合，其实在最近的十多年里已经不断涌现。图表 3 给出了一些人工智能在投资行业应用的案例。

图表 3、人工智能在投资领域的部分应用案例

时间	事件
2007 年	2007 年，纽约对冲基金 Rebellion Research 推出人工智能产品。
2008 年	国内首个大数据基金——百发 100
2008 年	首个智能投顾平台诞生：Wealthfront
2011 年	基于 Twitter 用户情绪变化的基金 Derwent Capital Markets 成立。
2013 年	2013 年，Kensho 公司成立并推出数据分析工具 Warren，公司于 2015 年被高盛以 1500 万美元收购。

2015 年	JP Morgan 的对冲基金业务 Highbridge Capital Management 开始与 Sentient 合作创建 AI 策略。Sentient 在资金方面获得 1.43 亿美元的支持，完全通过 AI 进行操作；在算法上，它使用了进化学习，生成了数万亿种专业程序
2017 年 6 月	华夏基金与微软公司在亚太地区设立的微软亚洲研究院举办战略合作发布会，宣布双方将就人工智能在金融服务领域的应用展开战略合作研究
2017 年 10 月	PINTEC（品钛）集团与香港富卫集团（FWD）合作在新加坡成立金融科技公司 PIVOT，这是中国智能投顾产业的首次出海。富卫集团为香港盈科拓展旗下的子公司，而盈科拓展的掌门人是李嘉诚之子李泽楷
2017 年 11 月	工商银行基于人工智能技术的智能投顾品牌“AI 投”已上线运行
2018 年 3 月	贝莱德正式推出了以“进化”统一命名的 7 只人工智能 ETF，这 7 只产品覆盖了消费、金融、医疗保健、科技、媒体等领域，是全球首个以人工智能推出的行业 ETF
2019 年 3 月	同花顺旗下人工智能资产管理有限公司同花顺阿尔法一号私募证券投资基金完成私募投资基金备案

资料来源：兴业证券经济与金融研究院整理

本文的主要研究对象，是机器学习方法在量化选股中的应用。在过往的量化股票投资实践中，多数的管理人倾向于使用简单明了的线性多因子框架来管理组合。究其原因，我们认为可能以下几个方面：首先，相对于线性模型来说，很多机器学习方法更像是一个黑盒，运算机制比较复杂，很多结果也不一定能找到相应的经济或金融学上的解释。其次，机器学习模型中包含的参数往往较多，这使得在训练建模的过程中，容易遭受过度优化问题的困扰，而这也会进一步导致模型的泛化能力下降，样本外预测能力不足。但复杂的非线性机制、灵活的参数选择本身就是机器学习方法的特点，也正是因为这种特点，才让机器学习模型在对金融市场运行规律的表达方面有着不同于传统线性模型的独特优势。我们认为，各种方法本身并无优劣之分，关键是看投资管理人如何去有效地运用。

本文的目标是基于机器学习模型，构建一个完整因子选股体系，以获取有别于传统线性模型的、更为有效的投资组合。

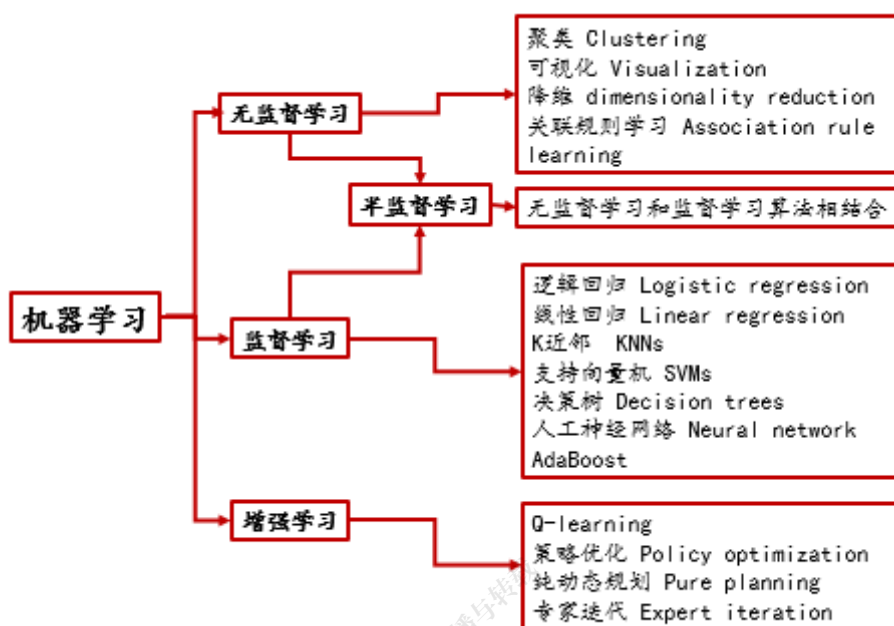
1.2、机器学习之集成学习

这里并不会对机器学习的整体做过多的阐述。我们将结合整个研究框架，有的放矢的对一些常见的概念做简单的介绍。如果欲详细了解机器学习，还请参考其他材料。

机器学习是计算机科学的一个分支，通过算法的设计，使得计算机具有一定的人工智能特性，以完成更为复杂的任务，例如：模式识别、系统控制、推断和预测等等。机器学习算法基于已获得的数据，深度挖掘其内在特征以及不同变量间的隐含关系，对可能存在的规律进行自动识别，并据此做出合理的判断。

机器学习主要有三种类型，分别为：监督式学习、无监督式学习以及增强学习。而常用的算法则包括：逻辑回归、决策树、回归树、支持向量机、神经网络、随机森林等。

图表 4、机器学习算法分类



资料来源：兴业证券经济与金融研究院整理

上面列出的种种算法，应该说各有优缺点，能否将不同的分类算法组合在一起，实现更好的学习分类效果呢？这就很自然地引出了本文研究的重点——集成学习方法（Ensemble Method）。

集成学习方法通过组合多个基学习器来完成学习任务，类似于我们俗语中说的“三个臭皮匠顶个诸葛亮”。基分类器一般采用的是弱学习器，通过集成方法，组合成一个强学习器。集成学习的泛化能力一般比单一的基学习器要好。集成学习方法主要包括 Bagging 和 Boosting 两种，Bagging 和 Boosting 都是将已有的分类或回归算法通过一定方式组合起来，形成一个性能更加强大的学习器。

Bagging 是自举汇聚法（bootstrap aggregating）的简称。Bagging 依赖于对训练数据所采用自举采样法（bootstrap sampling），即有放回地采样数据，主要做法为：（1）从原始样本集中抽取训练集。每轮从原始样本集中使用 Bootstrapping 的方法抽取 n 个训练样本。共进行 k 轮抽取，得到 k 个训练集。（ k 个训练集之间是相互独立的）（2）每次使用一个训练集得到一个模型， k 个训练集共得到 k 个模型。（3）对分类问题：将上步得到的 k 个模型采用投票的方式得到分类结果；对回归问题，计算上述模型的等权平均作为最后的结果。

Boosting 是一种与 Bagging 很类似的技术。其本质区别在于，Boosting 采用重赋权（re-weighting）法迭代地训练基学习器，每次在本轮的基学习器中强调上一轮基学习器中被错误分类的样本（增加错误分类样本的权重），最后在基学习器之间采用序列式的线性加权方式进行组合，得到最终的分类或回归结果。

Boosting 没有对样本进行重采样，而是对样本的分布进行了调整。

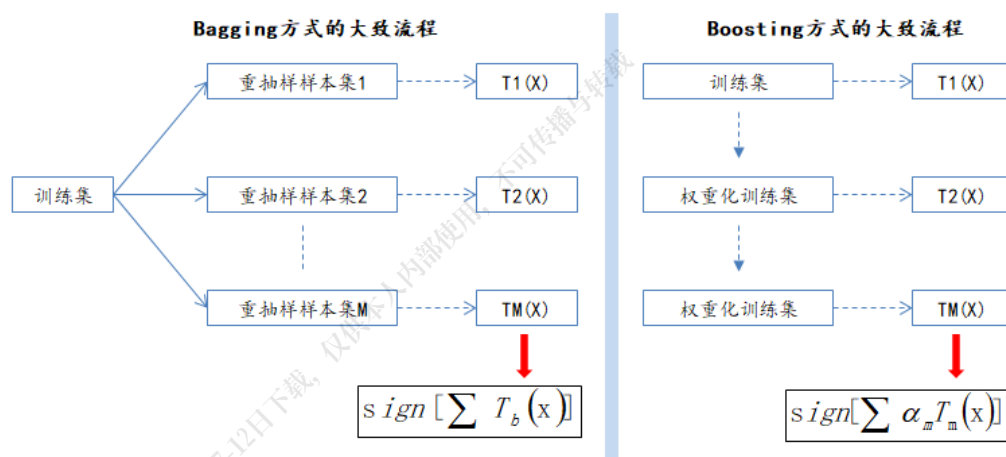
Bagging 与 Boosting 的区别具体体现在以下几个方面：

(1) 样本选择：Bagging 的训练集是在原始集中有放回选取的，从原始集中选出的各轮训练集之间是独立的。而 Boosting 每一轮的训练集不变，只是训练集中每个样例在学习器中的权重发生变化，权值是根据上一轮的分类结果进行调整。

(2) 样本权重：Bagging 使用均匀取样，每个样例的权重相等。Boosting 则根据错误率不断调整样例的权值，错误率越大则权重越大。

(3) 预测函数：Bagging 的所有预测函数的权重都是相等的。而 Boosting 下每个基学习器都有相应的权重，对于误差小的学习器会赋予更大的权重。

图表 5、Bagging Vs Boosting 流程



资料来源：兴业证券经济与金融研究院整理

集成学习的方法有很多，本文的研究对象 Adaboost 是 Boosting 方法中一类非常流行的算法，我们将深入探讨如何基于该算法构建一套非线性多因子选股模型的框架。后文中若无特别说明，研究的时间窗均为：2005 年 1 月 31 日-2019 年 4 月 30 日。

1.3、历史工作回顾

四年前，重剑出鞘，兴业金融工程团队耗时一年撰写了《聪明的 Alpha》系列报告，是市场上第一家系统性的将机器学习应用于 A 股市场量化选股领域的卖方团队。在这 5 篇深度报告里，我们详尽地探讨了以下几个方面的内容，包括机器学习的基本原理、应用于量化选股时需要注意的问题、不同参数设定下（如行业市值中性化的影响、因子池的优化、预测窗口的影响等）选股效果的差异、使用不同模型（Adaboost、Logistics、KNN、Random Forest、SVM、Decision Tree 等）选股效果的比较。文章研究角度新颖，覆盖议题广泛，在当时市场中收获了不错的口碑。

图表 6、过往研究回顾

报告	时间	主题
聪明的 Alpha 机器觉醒	2015/06/15	介绍了基于 Adaboost 算法构建的动态选股模型，并研究了日历效应
聪明的 Alpha 技术革命	2015/09/10	构建了仅仅利用技术因子&Adaboost 算法的动态选股模型，并研究了行业市值中性化的影响
聪明的 Alpha 重装上阵	2015/11/05	延长预测窗口，观察机器学习的作用，并研究了支持向量机算法的选股作用
聪明的 Alpha 万剑归宗(上篇)	2016/04/20	探索随机森林以及决策树在选股方面的应用
聪明的 Alpha 万剑归宗(下篇)	2016/04/20	探索逻辑回归、KNN 算法在选股方面的应用

资料来源：兴业证券经济与金融研究院整理

时隔四年，再次将目光聚焦于机器学习选股，我们已不是毫无准备。基于对过往研究的梳理和不断加深的认知，我们在继承之前成功经验的同时，也不断开拓创新，这主要体现在以下几个方面：

1. 因子池扩容：4 年前，我们量化因子库仅有 33 个选股因子，何其单薄。现如今，经过团队的不断积累，我们构建了基于价值、成长、质量、动量反转、分析师情绪、另类（细分为规模、风险、流动性、技术指标）维度的 165 个选股因子，多维度输入正好可以进一步发挥机器学习方法上的优势。
2. 模型算法改进：传统的 Adaboost 算法有一定的局限性，比如容易受极端值影响、分类结果较为粗糙等。此次我们使用的是改进 Adaboost 算法，其对选股问题的适用性更强，效果也更好，具体参见后续的介绍。
3. 日历效应再探：日历效应的本质是在特定的月份或者时期，市场的收益、风险表现明显异于其他时点。本报告在解决日历效应方面和之前的工作有一定的差异，设定更为合理。
4. 周度模型搭建：近年来，换手率较高的量化选股策略在业界越来越流行，有别于之前开发的月度模型，我们在本文中尝试利用价值类指标构建了相对高频的周度选股模型。此类模型对短期市场变化反应更为敏锐，策略收益率也有了显著提升。
5. 指数增强策略体系构建：对于目前的 A 股量化投资而言，指数增强是非常重要的且为机构投资广泛认可的一类产品形式。在本文的最后，我们利用构建的月度和周度选股模型，依据不同的风险收益偏好，设计了一套相对完整的指数增强策略跟踪体系。策略整体表现十分出色，后期我们也将定期跟踪样本外表现。

接下来，就让我们从基本概念出发，向大家展示基于 Adaboost 算法所构建的机器学习选股模型的每一个细节。Let's go!

2、基于 Boosting 集成学习的量化选股模型

2.1、Adaboost 算法介绍

➤ 传统 Adaboost 算法介绍

作为集成学习方法重要的一个子类，Boosting 算法在实践中有着广泛的应用，而 Adaboost 模型则是其最为流行的一种实现。传统的 Adaboost (Adaptive Boost

的缩写) 算法由 Yoav Freund 和 Robert Schapire 在 1995 年提出。在实现 Boosting 方法时, 有两个必须注意的问题:

- 1) 如何在基学习器中强调上一轮基学习器中被错误分类的样本 (损失函数的确定);
- 2) 在基学习器之间采用序列式的线性加权方式进行组合 (权重函数的使用)。

Adaboost 算法通过引入指数损失函数, 基于单层决策树实现了 Boosting 类的集成学习模型, 其在二分类问题背景下的算法步骤如下:

1. 模型基本设定: 训练样本集为 $\{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_i, y_i), \dots, (\mathbf{x}_N, y_N)\}$, 其中:

$y_i \in Y = \{-1, 1\}$, $\mathbf{x}_i = (x_{i1}, \dots, x_{ik}, \dots, x_{iK}) \in X$, x_{ik} 是第 i ($i=1, \dots, N$) 个样本在变量 k ($k=1, \dots, K$) 上的取值, N 为样本总数, K 为考察变量的总数 (后文中将用因子数来代替变量数), 样本的初始化权重向量为等权向量 $D_1(i) = \frac{1}{N}$ 。

2. 为了构建第 $l=1, \dots, L$ 层弱分类器:

- 1) 使用具有权重 D_l 的样本来训练弱分类器;
- 2) 可得到一系列的弱分类器 (一般为单层决策树) $g_k: X \rightarrow \{-1, 1\}$, 计算分类错误率 ε_k , 这里 $k=1, \dots, K$;
- 3) 遍历每一个 k , 得到使得 ε_k 最小的 $k^l \in \{1, \dots, K\}$, 那么第 l 层的弱分类器就确定为 $g_l \triangleq g_{k^l}$, 分类错误率定义为 $\varepsilon_l \triangleq \varepsilon_{k^l}$ 而该对该弱分类器的信心度则使用下式来定义:

$$h_l = \frac{1}{2} \ln \left(\frac{1 - \varepsilon_l}{\varepsilon_l} \right) \quad (1)$$

不难发现, 错误率 ε_l 越小, 弱分类器的信心度也就越高, 对最后的强分类器的影响也就越大;

- 4) 更新样本权重向量

$$D_{l+1}(i) = \frac{D_l(i)}{Z_l} \exp(-h_l y_i g_l), \quad (2)$$

这里 Z_l 为归一化因子。

3. 强分类器的生成: 最终的强分类器, 将由每一层确定的弱分类器和相应的信心度共同确定得到:

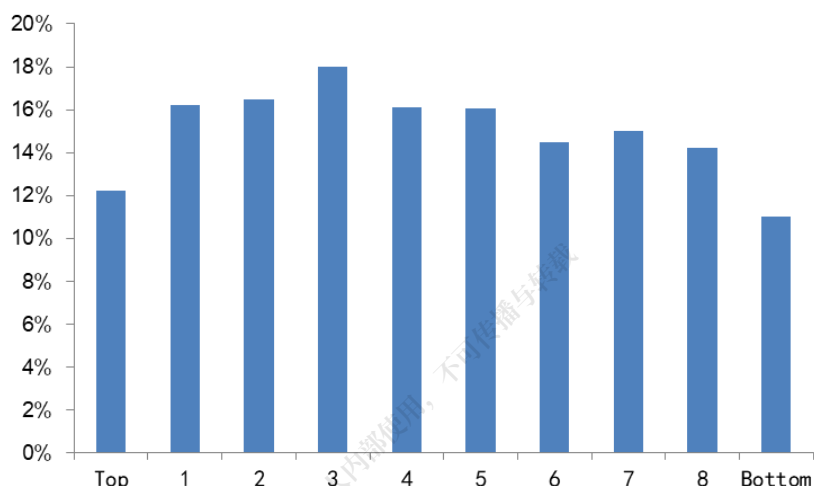
$$G(\mathbf{x}) = \text{sign} \left(\sum_{l=1}^L h_l g_l(\mathbf{x}) \right) \quad (3)$$

➤ 改进 Adaboost 算法介绍

上述的传统算法在很多问题上都取得了良好的表现, 但当我们考虑将 Adaboost 应用于量化选股问题时, 其存在的一些问题也不可避免的暴露了出来。首先, 量化选股往往体现为一个排序问题, 但传统二分类算法的最终输出结果仅

为 1 或-1，在生成最终股票排名时不太方便，区分度不够；其次，在构建弱分类器时，传统方法本质上还是假设了因子与最终分类结果之间满足单调性，这对于线性分类问题没什么太大的影响，不过一旦遇到因子与结果之间呈现非线性相关的情况，传统 Adaboost 算法就力不从心了。下面的图表 7，展示了我们在《猎金系列之二十八——公司治理与股票预期收益率》中构建的“股权集中度”因子与预期收益率之间的关系，可以发现两者呈现一种倒 U 型相关性。可以想象，如果仅采用传统 Adaboost 算法，我们就无法充分利用股权集中度因子进行选股。

图表 7、股权集中度因子分位数测试效果



资料来源：Wind，兴业证券经济与金融研究院整理

为了应对上述问题，我们引入了改进版的 Adaboost 模型，具体算法流程如下：

- 模型基本设定：训练样本集为 $\{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_i, y_i), \dots, (\mathbf{x}_N, y_N)\}$ ，其中：
 $y_i \in Y = \{-1, 1\}$ ， $\mathbf{x}_i = (x_{i1}, \dots, x_{ik}, \dots, x_{iK}) \in X$ ， x_{ik} 是第 i ($i=1, \dots, N$) 个样本在变量 k ($k=1, \dots, K$) 上的取值， N 为样本总数， K 为考察变量的总数（后文中将用因子数来代替变量数），样本的初始化权重向量为等权向量 $D_1(i) = \frac{1}{N}$ 。
- 为了构建第 $l=1, \dots, L$ 层弱分类器：
 - 使用具有权重 D_l 的样本来训练弱分类器，对于 $\forall k \in \{1, \dots, K\}$ ，将训练样本 $\{(x_{1k}, y_1), (x_{2k}, y_2), \dots, (x_{Nk}, y_N)\}$ 按照因子值 $\{x_{1k}, x_{2k}, \dots, x_{Nk}\}$ 的大小等分为 Q 组： $X_1, \dots, X_Q$ ，其中 X_Q 中包含了 $\{x_{1k}, x_{2k}, \dots, x_{Nk}\}$ 中取值最大的 $\left\lceil \frac{N}{Q} \right\rceil$ 个样本（这里 $\lceil \cdot \rceil$ 代表向上取整）。
 - 计算 Q 组样本中，每一组内 y_i 取值为 1 和-1 的样本的权重和：

$$W_k^{q\pm} = \sum_{y_i = \pm 1, x_{ik} \in X_q} D_l(i), q=1, \dots, Q, \quad (4)$$

这里不难发现 $W_k^{q+} + W_k^{q-} = 1$ ，定义错误率指标为

请务必阅读正文之后的信息披露和重要声明

$$\varepsilon_{lk} = \sum_{q=1}^Q \sqrt{W_k^{q+} W_k^{q-}}, \quad (5)$$

仔细观察上式不难发现,在任意分组 q 内,1与-1样本的数量差距越大(即分类效果越好),错误率 ε_{lk} 的取值就越小;反之,分类效果越差, ε_{lk} 取值也会越大;

- 3) 遍历每一个 k ,得到使得 ε_{lk} 最小的 $k^l \in \{1, \dots, K\}$,定义第 l 层弱分类器的信心度为:

$$h_l^q = \frac{1}{2} \ln \left(\frac{1 - W_{k^l}^{q-}}{W_{k^l}^{q-}} \right) = \frac{1}{2} \ln \left(\frac{W_{k^l}^{q+}}{W_{k^l}^{q-}} \right) \quad (6)$$

注意,这里的信心度是一个带符号的数值,当 $W_k^{q+} > W_k^{q-}$ 时,信心度取值为正,且差距越远, h_l^q 的取值越大,表明分类结果更可能为1;当 $W_k^{q+} < W_k^{q-}$ 时,信心度取值为负,且差距越远, h_l^q 的取值越小,表明分类结果更可能为-1;当 $W_k^{q+} = W_k^{q-}$ 时,信心度取值为0,表明模型对该样本没有区分能力。

- 5) 更新样本权重向量

$$D_{l+1}(i) = \frac{D_l(i)}{Z_l} \exp(-h_l^q y_i), \quad (7)$$

这里 Z_l 为归一化因子。

4. 强分类器的生成: 最终的强分类器,将由每一层确定的弱分类器的信心度相加得到:

$$G = \sum_{l=1}^L h_l^q \quad (8)$$

将传统 Adaboost 与其改进版进行对比后不难发现,改进版本在以下几个维度做出了调整:

- 1) 有别于传统算法中对因子值直接二分的做法,改进算法中引入了 Q 个分位组合,这使模型对非线性关系的表达得到了加强,分类的精细度有所提升;
- 2) 针对改进后的做法重新设计了错误率的定义方式(见公式(5)),能够更稳健地在多分组的情况下对模型的整体区分度进行衡量;
- 3) 直接利用弱分类器信心度之和反映最终的分类结果(见公式(8))。

2.2、基于改进 Adaboost 算法的量化选股模型

在上一节中我们指出,改进版的 Adaboost 算法从特点上讲更有利于量化选股模型的表达。在本节中我们将详细阐述基于改进 Adaboost 算法的量化选股框架。总的来说,我们将量化选股问题化归为了一个二元分类问题,根据下期股票收益

率的大小，将股票池划分强势股和弱势股两类。模型的作用就是要给出某只股票属于强势股还是弱势股的分类结果。结合改进版 Adaboost 算法，模型具体的实现流程参见图表 8。

图表 8、改进 Adaboost 算法构建选股模型流程图



资料来源：兴业证券经济与金融研究院整理

模型的输入是我们兴业金融工程团队构建的包含了 165 个量化选股指标的因子库。该数据库包括了价值、成长、质量、分析师情绪、动量反转、另类这六大类指标，其中另类进一步分为：规模、风险、流动性、技术这四个子类（部分因子定义参见附录）。最终输出的结果是对每个股票的信心度，用于衡量该股票属于强势股的程度（取值越大，越有可能属于强势股；反之则越有可能属于弱势股）。

图表 9、兴业量化选股因子分类数量统计

	价值	成长	质量	分析师情绪	动量反转	另类
因子数	19	33	39	13	9	52

资料来源：Wind，兴业证券经济与金融研究院整理

➤ 训练样本

首先我们要进行训练样本的准备。大多数的机器学习算法对于输入的数据都较为敏感。如果训练数据噪声较多，一致性不好，那么算法极有可能陷入过度优化的泥沼。所以说数据样本准备对于机器学习来说，是至关重要的一步。在某种意义上，甚至超过了对学习算法的选择。

我们使用横截面上**因子的序**作为输入，而不是简单使用原始因子值，这在某种程度上也能规避极端值对算法的影响。同时因为我们的目的是比较股票的相对优劣，而并非关注具体数值的大小，因子的序完全满足这一要求。在每个时间点上，对于每个因子，我们计算每只股票按照这个因子排列的顺序，并除以总的股票数，使得所有因子的值都划归到 (0,1] 之间。

对于收益率数据我们也要做一些处理。我们将下期股票收益率从大到小排列，

请务必阅读正文之后的信息披露和重要声明

取前 30% 的股票作为强势股，后 30% 作为弱势股，中间的股票视作噪声弃之不用。强势股我们标记为+1，弱势股我们标记为-1。这样每个样本点可以记作 $(\mathbf{x}, y)$ ，其中 $\mathbf{x} = (x_1, \dots, x_k, \dots, x_K)$ 是股票在各个因子上的序化的暴露， $y = \pm 1$ 。

为保证算法的稳定性，我们用过去 12 个月月底的面板数据来构建训练样本。

➤ 弱分类器

每个弱分类器可以简单的理解为一个因子。对于每个因子，我们按照其分位点将股票等数量划分为若干组，每一组内我们分别计算强势股和弱势股的权重总和。直观上讲，区分能力最有效的因子应该是强势股和弱势股权重和相差较大的因子。遍历所有的因子，我们选择区分能力最强的因子作为该层的弱分类器，同时利用该因子每个分位组里强势股和弱势股权重孰大作为区分的标志。当新的因子值落入某个分位组里时，如果该分位数组里强势股权重大，我们便将该股票分类为强势股，反之则为弱势股。

➤ 权重更新

初始时样本权重为等权，在每次找到新的弱分类器后，我们将调整数据的权重。被正确区分的样本权重将被调低，而区分错误的样本权重将加大。下一层的分类器将基于调整后的权重进行构建。这样每次我们都将学习的重点转移到上一层分类错误的股票上，最终将逐步改善模型整体的分类准确度。

➤ 强分类器

最后，我们将所有弱分类器加总得到强分类器。分类器会赋给每一组内的股票一个数值作为信心度，用以衡量其属于强势股的程度，值越大越可能为强势股，越小越可能为弱势股。我们把这个强分类器输出的信心度分值称为 **ELS 因子** (Ensemble Learning Signal, 即集成学习因子)。

上述四个步骤的具体实现可参考 2.1 节中对改进版 Adaboost 算法的描述。在后文分析参数敏感性之前，我们取定分组数 $Q=5$ ，弱分类器的总层数 $L=30$ 。由于我们样本的起始日期为 2005 年 1 月底，考虑到模型需要滚动 12 个月的样本作为训练集，输出的 ELS 因子从 2006 年 1 月底可以开始计算。

2.3、ELS 因子表现

从因子测试的表现来看：ELS 因子 IC 均值达到 0.098，年化 IC_IR 高达 3.22，t 值为 11.8，有效性非常显著；观察因子的分位数表现，各组夏普率、年化收益率严格单调，多空组合年化收益率达到 39%，夏普比率达到 3.39，但最大回撤比较明显，为 16.4%。

图表 10、ELS 因子 IC 测试结果

	平均值	标准差	最小值	最大值	IC_IR	t 统计量
ELS	0.098	0.11	-0.29	0.40	0.93	11.8

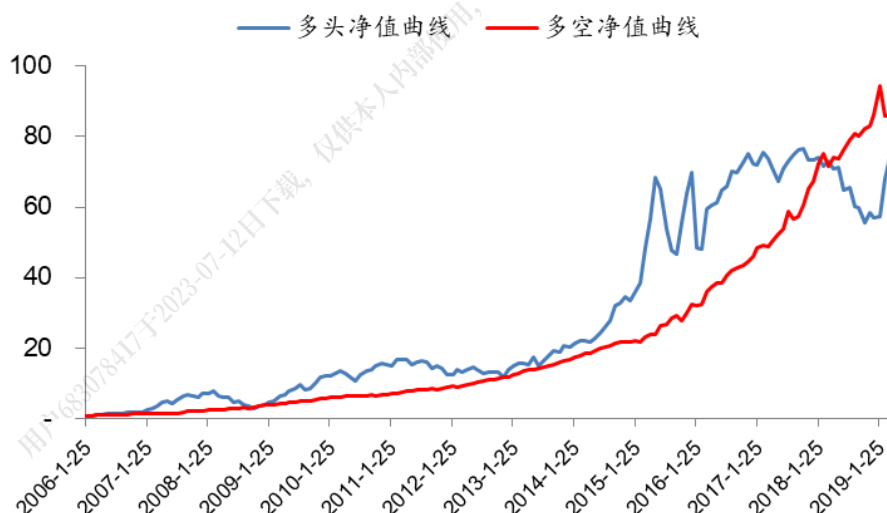
资料来源：Wind，兴业证券经济与金融研究院整理

图表 11、ELS 因子分位数组测试效果

	总收益率	年化收益率	Sharpe 比率	换手率	最大回撤率	年化超额收益率	跟踪误差	信息比率	胜率
top	7284.3%	38.4%	1.22	6.0%	64.9%	14.6%	5.9%	2.49	57.1%
1	4247.3%	33.0%	1.07	7.9%	65.7%	10.0%	4.6%	2.18	55.0%
2	2677.7%	28.5%	0.93	8.3%	70.0%	6.3%	3.7%	1.70	53.8%
3	2188.5%	26.7%	0.88	8.5%	69.7%	4.6%	3.8%	1.22	53.0%
4	1301.2%	22.1%	0.73	8.5%	70.1%	0.8%	3.2%	0.25	50.4%
5	1343.4%	22.3%	0.75	8.6%	69.2%	0.9%	3.3%	0.27	48.8%
6	906.7%	19.1%	0.63	8.5%	70.7%	-1.7%	3.5%	-0.48	48.0%
7	450.6%	13.8%	0.45	8.3%	73.5%	-6.0%	3.9%	-1.54	45.0%
8	259.4%	10.1%	0.32	8.0%	73.2%	-8.7%	4.8%	-1.79	44.2%
bottom	-19.6%	-1.6%	-0.05	6.5%	84.1%	-18.1%	7.1%	-2.55	41.4%
L_S	1134.1%	20.9%	0.69		70.4%				
市场	7691.5%	39.0%	3.39		16.4%				

资料来源：Wind，兴业证券经济与金融研究院整理

图表 12、ELS 因子多头以及多空组合表现



资料来源：Wind，兴业证券经济与金融研究院整理

2.4、行业市值中性化

考虑到历史上行业 and 市值因子对 A 股横截面收益率预测的重要意义，我们尝试分析行业市值中性化对 ELS 因子的影响，观察在剔除市值或行业风险后，是否对目标因子风险收益比的提升有所帮助。

在进行行业市值中性化的处理时，我们采取如下的方法：在每个行业内，以行业内所有股票的流通市值中位数为界来进行大小票的划分，中位数以上者为该行业的大盘股，以下者视为该行业小盘股。而后分别在每个行业市值股票池内进行横截面因子的 Rank 标准化，再输入到模型中进行训练。通过这种方式所构建的

ELS 因子，称为 NELS (Neutralized Ensemble Learning Signal) 因子。

我们对 NELS 因子也进行了全面的测试：从 IC 测试结果来看，中性化之后的因子的 IC、IC_IR 以及 t 值要远优于未作中性化处理的因子表现；同时 NELS 因子过去 6 个月的 IC 均值走势明显要更为平稳。而分位数测试的结果表明，NELS 要全面优于 ELS，多空组合的年化收益率、夏普率分别达到 55.3%、5.6，从收益和风险两个角度都提供了更好的表现。

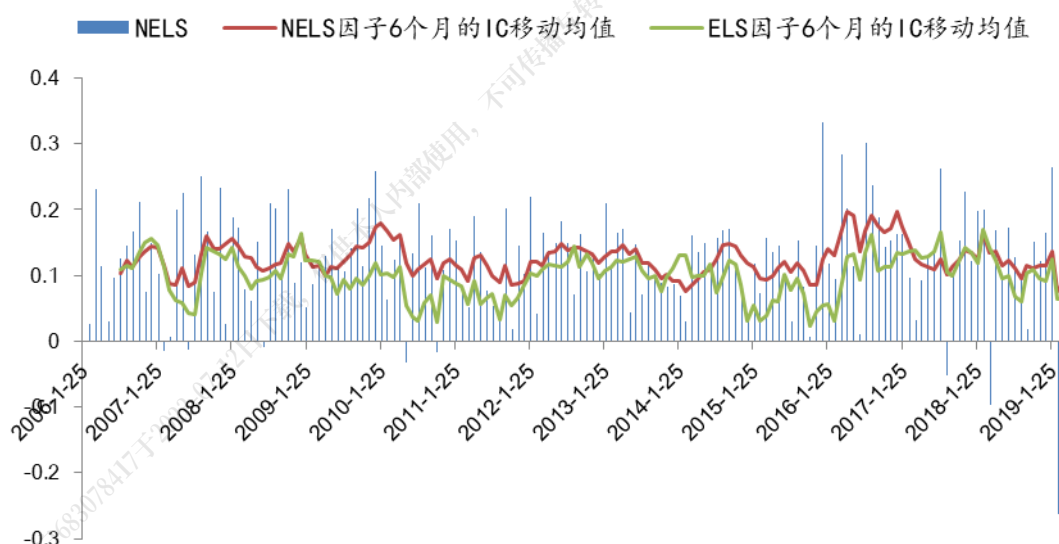
综上，我们后续的研究将以行业市值中性化之后的结果作为研究对象。

图表 13、NELS Vs. ELS 因子 IC 测试

	平均值	标准差	最小值	最大值	IC_IR	t 统计量
ELS	0.098	0.11	-0.29	0.40	0.93	11.8
NELS	0.123	0.08	-0.26	0.33	1.59	20.0

资料来源：Wind，兴业证券经济与金融研究院整理

图表 14、NELS Vs. ELS 因子 IC 历史走势



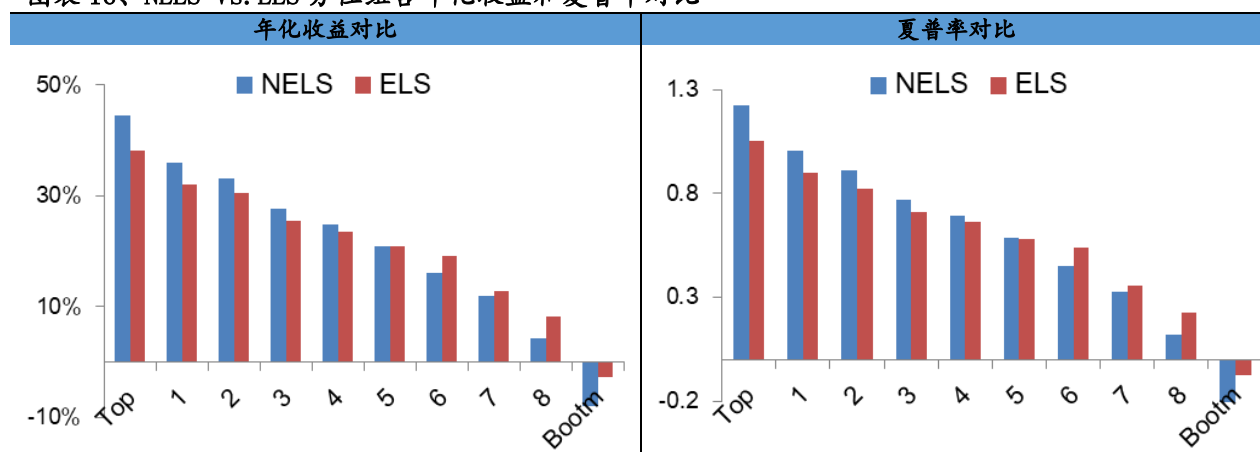
资料来源：Wind，兴业证券经济与金融研究院整理

图表 15、NELS Vs. ELS 分位数组合测试结果

	NELS 因子表现				ELS 因子表现			
	年化收益	Sharpe	换手率	信息比	年化收益	Sharpe	换手率	信息比率
top	44.5%	1.22	143.8%	3.90	38.2%	1.06	127.5%	2.55
1	35.9%	1.00	167.0%	3.50	32.1%	0.90	161.1%	2.20
2	33.0%	0.91	172.5%	2.83	30.5%	0.83	168.6%	1.91
3	27.6%	0.77	174.5%	1.79	25.4%	0.71	170.5%	1.37
4	24.7%	0.69	174.3%	1.09	23.5%	0.66	171.4%	0.70
5	21.0%	0.59	174.7%	0.10	20.8%	0.58	172.0%	0.03
6	16.1%	0.45	174.5%	-1.16	19.1%	0.54	171.6%	-0.44
7	11.9%	0.33	172.7%	-1.91	12.7%	0.35	169.9%	-1.77
8	4.3%	0.12	168.2%	-2.83	8.3%	0.23	162.8%	-2.15
bottom	-8.1%	-0.22	149.0%	-3.83	-2.8%	-0.08	135.0%	-2.93
市场	20.5%	0.57			20.5%	0.57		
L-S	55.3%	5.60			40.5%	3.85		

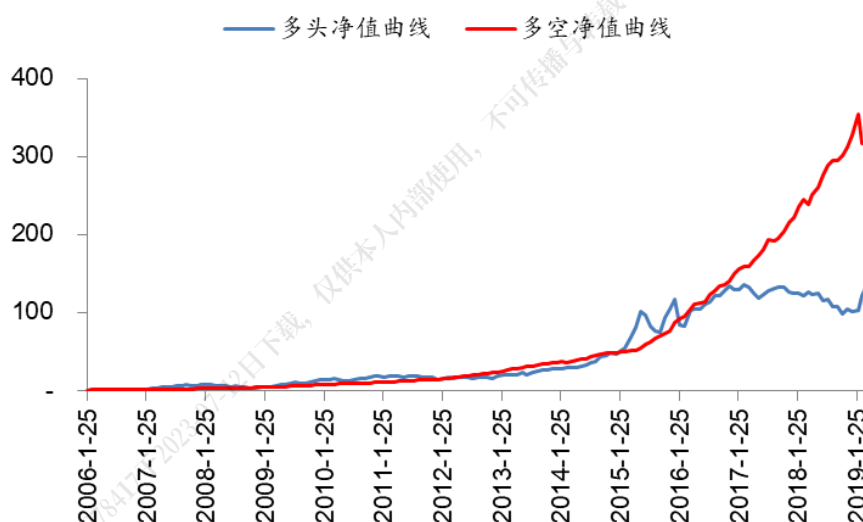
资料来源：Wind，兴业证券经济与金融研究院整理

图表 16、NELS Vs. ELS 分位组合年化收益和夏普率对比



资料来源：Wind，兴业证券经济与金融研究院整理

图表 17、NELS 因子的多空净值曲线



资料来源：Wind，兴业证券经济与金融研究院整理

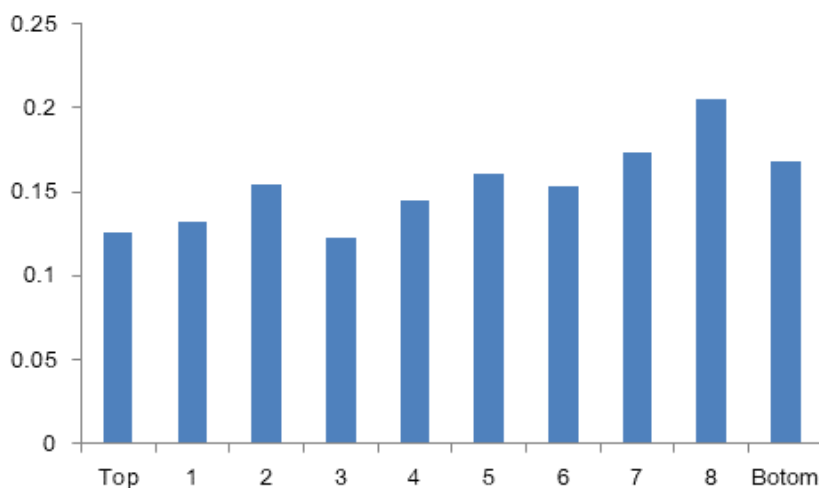
2.5、模型优势

基于改进版 Adaboost 算法的量化选股模型具备一系列独特的优势，下面我们将逐一展开讨论。

1. 非线性因子的利用

在引入改进 Adaboost 算法时，我们提到：该算法在截面上将所有的股票分成 Q 组，然后对每一组内强势股与弱势股进行区分。这种细分方法一定程度上将非线性关系转化成了线性关系（比如选股效果为 U 型的因子）。仍然以我们前面提及的管理层集中度指标为例，经过 Adaboost 算法调整之后的因子分位数组合呈现一定程度的单调性，选股效果大幅优于之前的表现（具体参见图表 7）。

图表 18、修正后的股权集中度因子分位数测试结果



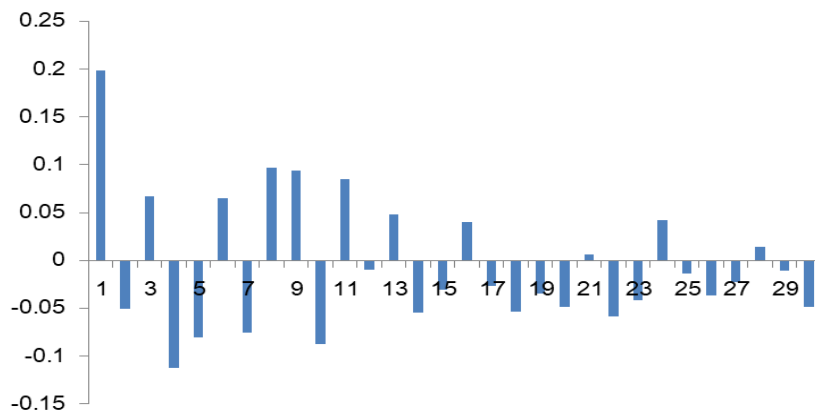
资料来源：Wind，兴业证券经济与金融研究院整理

2. 避免过度拟合

对于机器学习而言，过拟合与欠拟合都是我们需要解决的问题。如果样本内的训练表现过好，那么样本外的泛化能力必然会受到影响；如果样本内的训练表现过差，也很难相信该模型能够对样本外的判断起到什么作用。而 Adaboost 算法天然对过拟合问题具有免疫性，这与其独特的算法实现流程不无关系：首先在每一层弱分类器的选择过程中，我们总是选择增量信息最大（注意是增量信息，这意味着和已选择的弱分类器的互补性）的指标以保证效果的提升是最优的；其次，后入选指标的重要性几乎是以指数形式衰减的，速度非常快，这也就意味着即便多增加几个弱分类器，其对最终结果的影响也不大。

为了更直观的观察因子的重要性，我们将 2019 年 1 月 31 日的训练结果单独提取了出来，并绘制了图表 19。从图中不难发现，首先入选的几个因子贡献了绝大多数话语权，而后续的因子边际贡献实际上是很小的。

图表 19、因子重要性程度示意图



注：横坐标代表入选弱分类器的序，纵坐标代表五组中第一组因子输出大小

资料来源：Wind，兴业证券经济与金融研究院整理

请务必阅读正文之后的信息披露和重要声明

3. 规避多重共线性

对于困扰许多量化选股模型的多重共线性问题，Adaboost 算法也能有效规避，因为从原理上来讲，Adaboost 算法始终寻求增量效应最大化，一旦某一类相关性较高的因子中有一个成为了弱分类器，其他因子就很难再被选中。我们同样以 2019 年 1 月份选中的 30 个弱分类器为例进行说明。这一期选中的 30 个弱分类器中包含了 EP_FY1（一致预期净利润_FY1 / 总市值）。在 165 个因子中，和该因子相关性比较高的因子多达 8 个，但这 8 个因子却都没有进入到最终的弱分类器当中，在一定程度上印证了 Adaboost 算法在规避多重共线性上的能力。

图表 20、与 EP_FY1 相关性高的因子展示

因子	定义	相关性
EP_LYR	净利润(不含少数股东损益)_最新年报 / 总市值	0.661
EP_TTM	净利润(不含少数股东损益)_TTM / 总市值	0.708
EP_FY0	一致预期净利润_FY0 / 总市值	0.912
EP_Fwd12M	每股收益_未来 12 个月预测值 / 收盘价	0.983
EP_TTM_Deducted	扣除非经常性损益后净利润_TTM / 总市值	0.677
EP_LYR_Deducted	扣除非经常性损益后净利润_最新年报 / 总市值	0.626
EBITDA2EV	息税折旧摊销前利润_TTM / (总市值 + 带息债务_最新财报 - 货币资金_最新财报)	0.590
EBIT2EV	息税前利润_TTM / (总市值 + 带息债务_最新财报 - 货币资金_最新财报)	0.619

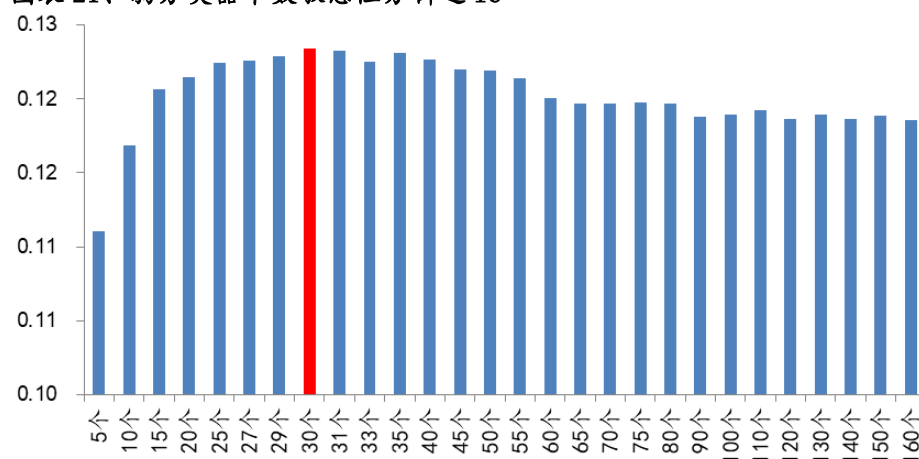
资料来源：Wind，兴业证券经济与金融研究院整理

4. 参数敏感性较弱

机器学习的参数敏感性问题是很多人关注的焦点。在本文应用的模型算法中，弱分类器的个数是一个主要参数。在前面的章节中，我们从 165 个因子池中选择了 30 个因子作为弱分类器。这里，我们将对弱分类器个数做敏感性分析，以确定合理的参数取值。

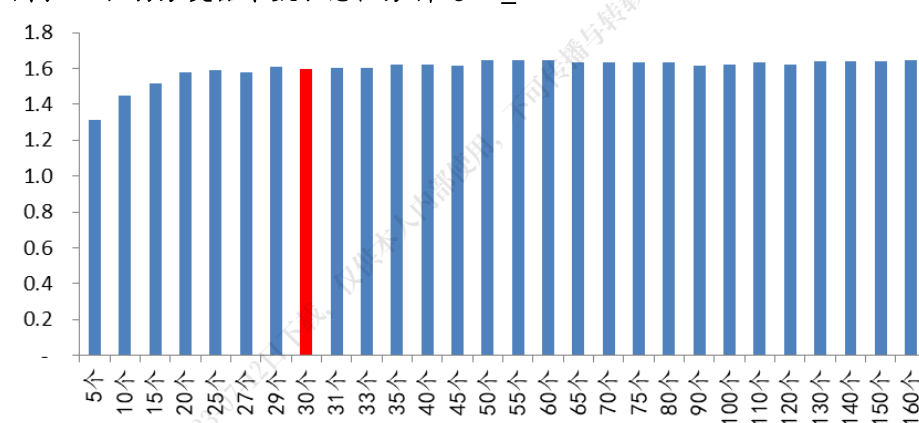
我们遍历了参数取值为 5、10、15……160 等的各种情况，从 IC 测试结果来看：当弱分类器数量由 5 个提升至 25 个时，边际改善非常明显；从 25-30 阶段也有一定的效果，而之后提升效果甚微，表现则趋于稳定；从因子的分位数测试结果来看，随着弱分类器的数目提升，多空组合的年化收益率以及夏普率的改善效果非常显著，当弱分类器数到达 30 个以后，边际改善效应明显递减。综上，我们将弱分类器个数定为 30。

图表 21、弱分类器个数敏感性分析之 IC



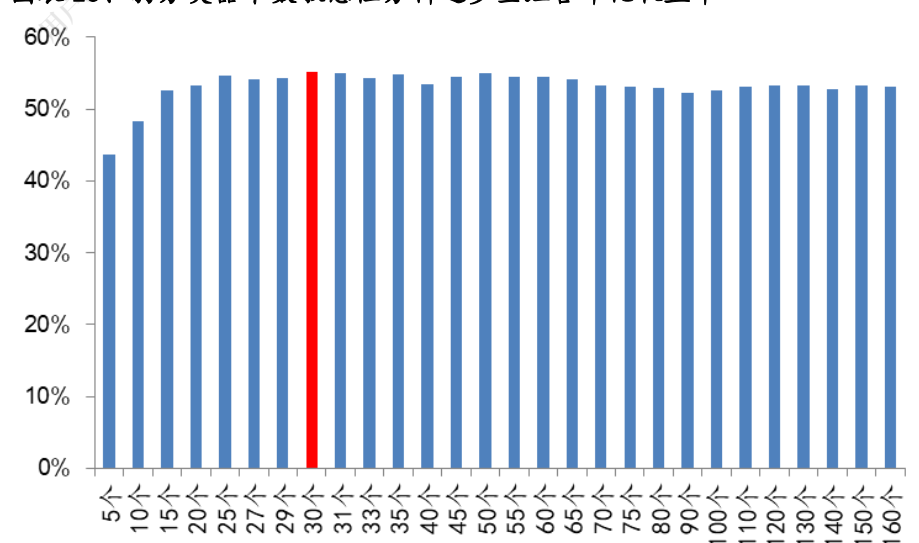
资料来源：Wind，兴业证券经济与金融研究院整理

图表 22、弱分类器个数敏感性分析之 IC_IR



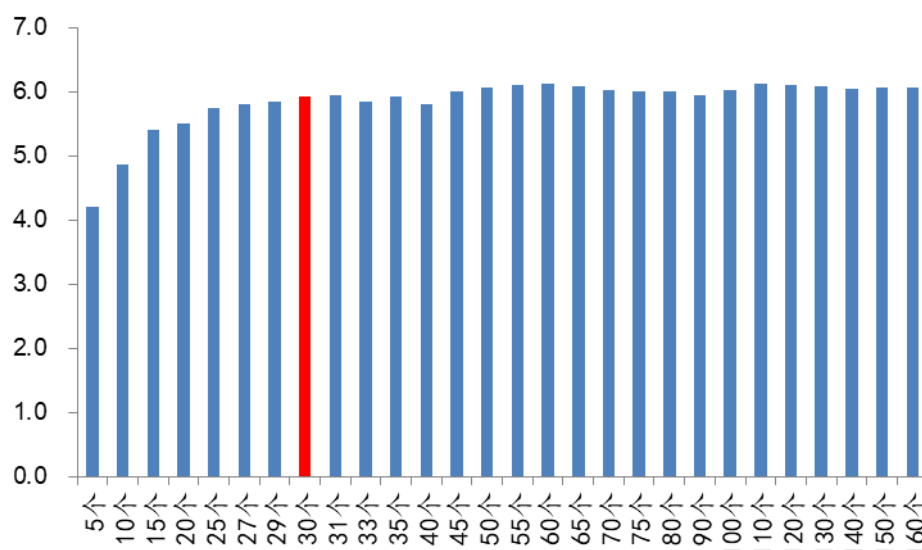
资料来源：Wind，兴业证券经济与金融研究院整理

图表 23、弱分类器个数敏感性分析之多空组合年化收益率



资料来源：Wind，兴业证券经济与金融研究院整理

图表 24、弱分类器个数敏感性分析之多空组合夏普率



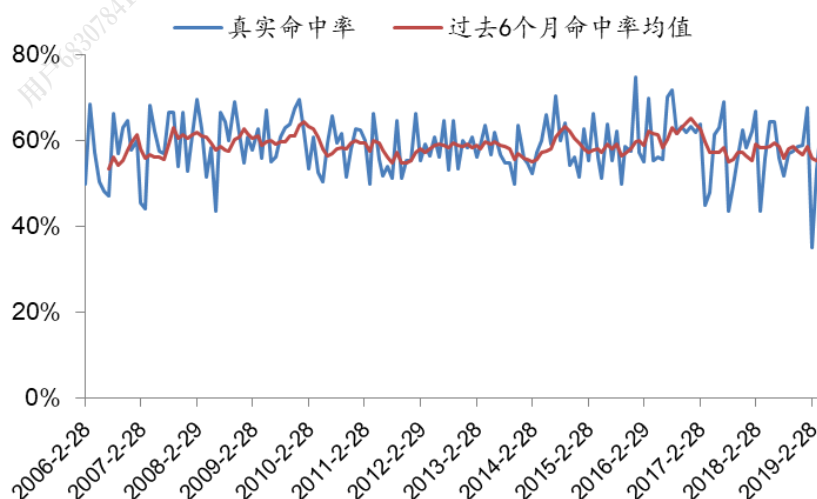
资料来源：Wind，兴业证券经济与金融研究院整理

5. 选股命中率较高

量化和主动研究的本质区别在于：量化推的是股票池、最终体现的是概率（推荐的股票池中有多少是上涨的），而主动推荐的是个股，结果是0或者1（要么上涨要么下跌）。量化模型一般会尽可能提升这种推荐命中率。

我们这里以分位数测试的 Top 组合为例展开研究，研究其命中率，即每期分位数首组选中的股票跑赢市场基准的概率。从测试结果来看，基于 NELS 因子的命中率平均可达 60% 左右。从历史上每一期的胜率来看，整体也比较稳定。这进一步说明了模型的有效性和稳定性。

图表 25、NELS 因子的选股命中率



资料来源：Wind，兴业证券经济与金融研究院整理

6. 动态模型 vs 静态模型

我们在上文中构建的基于改进版 Adaboost 算法的量化选股模型，从本质上讲

属于一类动态选股模型。相比于传统的权重（等权或其他加权方式）和因子（基于长期历史样本筛选得到）相对固定的静态模型而言，其优势主要体现在以下几个方面：

1. 静态模型多采用多元线性回归作为基本形态，对非线性规律把握难度较大，而且其因子和权重设置相对固定，很难适应国内市场规模快速切换的现实情况；而动态模型则会根据考察时间窗内的市场动态进行因子和权重的调整，自适应能力更强，更为灵活和有效。
2. 静态模型的搭建需要大量的历史数据做回测以大概率地保证因子的长期有效性，而动态模型的时变性特征仅要求我们去把握最近一段时间的因子规律，从而对历史样本的需求更少，而样本外测试结果的可信度也相对更高。

2.6、玩转风格轮动

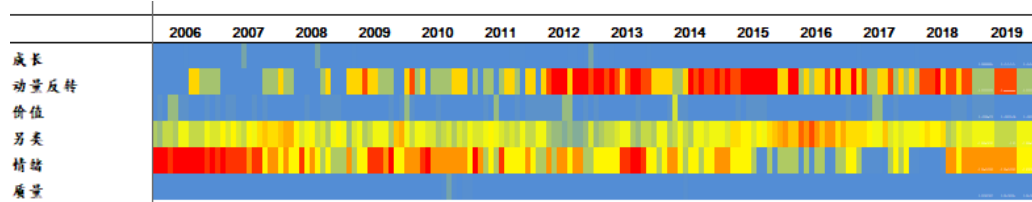
机器学习在量化选股中难以推广的一个原因是大家普遍认为其黑箱操作的性质，很难理解每次选股做了什么调整，选中了什么风格。而我们前面也提到了 Adaboost 算法流程相对容易理解，黑盒属性较弱。这里我们将历史上每期选择的因子提取出来并分析其风格。

由于不同风格本身因子池大小不一，如果仅仅看因子数目的话有失偏颇，所以我们这里会分析：1、每期排名前三的因子所属风格（即最为重要的三个弱分类器）；2、选中的不同风格的因子（选中的 30 个弱分类器）占本身因子池数目比例。

分析每期选中的 30 个弱分类器因子所属分类占相应风格的比例，我们发现动量反转、情绪、另类因子一直居于较高的水平，说明这些类别中的因子持续有效性较强。

从单个因子维度来看，另类因子有效性一直较强，但这里需要注意的是：Amount_20D_Avg、Turnover_20D_Avg 两个因子的表现与众不同：有效性最强、持续性好，并且形成了很好的互补性。Amount_20D_Avg 表示过去 20 天的日均成交量，在 08 年金融危机之后以及 2013-2016 年慢牛+牛市期间最为有效；Turnover_20D_Avg 表示过去 20 天的日均换手率，在 2008 年金融危机之前以及期间、2017-2018 的震荡和当下的市场最为有效。动量反转因子在 2011-2013 年期间表现较为强势，在这期间成长因子也表现出一定的优势；基于财务类的价值因子在 2006-2008 年以及 2016-2017 年表现出一定的优势，近期呈现复苏的迹象。

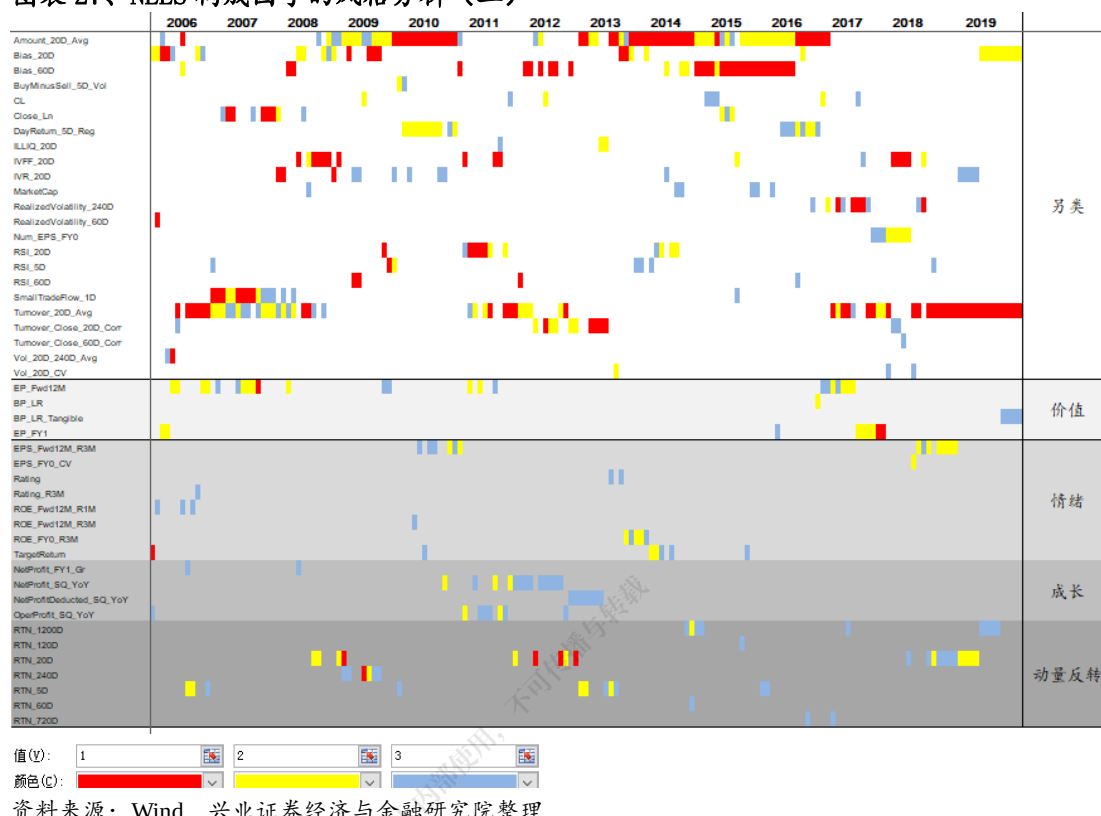
图表 26、NELS 构成因子的风格分析（一）



资料来源：Wind，兴业证券经济与金融研究院整理

注：色阶释义 [0, 0.15] [0.15, 0.3] [0.3, 1]

图表 27、NELS 构成因子的风格分析（二）



3、模型改进与增强

3.1、失败乃成功之母

在上一章节中，我们详细介绍了基于改进版 Adaboost 算法的量化选股模型，并给出了相应的 NELS 因子的表现。该阿尔法模型在历史上的预测能力是非常稳健的，但也出现过若干次明显的回撤和失效的情形。如何进一步提升模型的表现呢？加入更为有效的选股因子当然是一种方式，但在本节中我们要引入的方法其实更为直接，用一句俗语来概括，那就是：失败乃成功之母！

具体来说，站在某个时点，首先我们计算之前构建好的 NELS 因子的历史 IC 序列，然后将那些 IC 值低于历史 IC 中位数的时点的样本拼在一起，称之为“失败者样本”（之所以叫做“失败者样本”，是因为 NELS 因子在这些样本的中预测能力表现相对较差）。最后，我们在失败者样本上重新训练之前的选股模型，并利用得到的新因子 L-NELS（Loss NELS）修正原有的 NELS 因子。在实践中，我们要求 NELS 因子需要具有 10 年的历史数据，如果不足 10 年，至少要保证 80% 的非空率；因而本章节因子的计算和测试只能从 2013 年开始。

我们对基于失败者样本得到的 L-NELS 因子进行了全面的测试。从结果来看，因子 IC 达到 0.124、IC_IR 高达 1.49，分位数测试各组别严格单调，多空组合年化收益率 57.7%，夏普率达到 4.74，有效性非常强。

图表 28、L-NELS 因子的 IC 检验

	平均值	标准差	最小值	最大值	IC_IR	t 统计量
L-NELS	0.124	0.08	-0.05	0.35	1.49	12.8

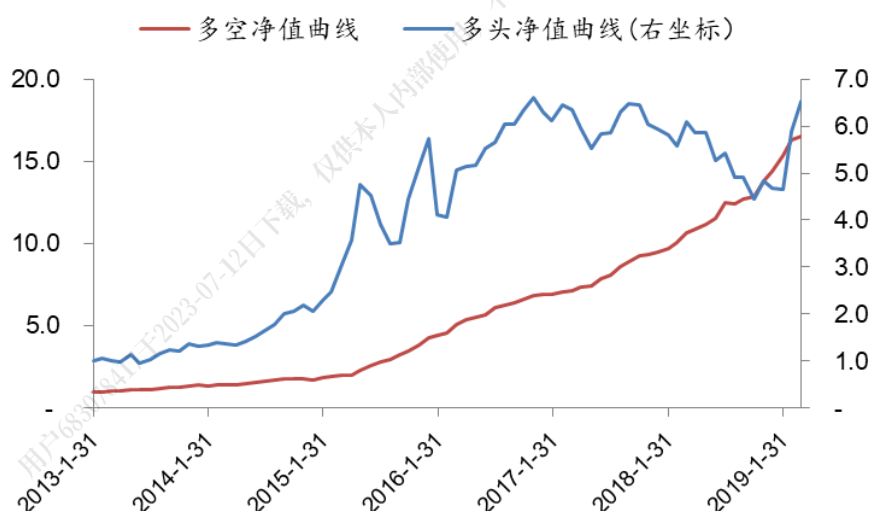
资料来源：Wind，兴业证券经济与金融研究院整理

图表 29、L-NELS 因子分位数组合测试结果

	总收益率	年化收益率	Sharpe 比率	换手率	最大回撤率	年化超额收益率	跟踪误差	信息比率	胜率
top	552.8%	35.6%	1.03	139%	32.8%	18.8%	6.5%	2.89	77.3%
1	430.0%	31.1%	0.92	166%	37.6%	14.5%	4.6%	3.17	85.3%
2	345.2%	27.4%	0.84	172%	36.6%	11.0%	3.7%	2.99	86.7%
3	285.0%	24.5%	0.78	173%	36.5%	8.1%	2.5%	3.18	81.3%
4	230.7%	21.4%	0.69	175%	42.6%	5.3%	3.2%	1.65	68.0%
5	182.1%	18.3%	0.60	174%	48.3%	2.4%	3.3%	0.74	57.3%
6	82.0%	10.2%	0.33	174%	59.1%	-4.7%	3.3%	-1.44	30.7%
7	56.2%	7.5%	0.25	172%	64.9%	-7.1%	3.5%	-2.02	16.0%
8	-7.5%	-1.3%	-0.04	166%	74.4%	-14.7%	4.9%	-3.02	12.0%
bottom	-62.0%	-14.5%	-0.46	142%	86.3%	-26.1%	6.8%	-3.82	1.3%
市场	139.2%	15.2%	0.48		52.2%				
L_S	1555.3%	57.7%	4.74		3.4%				

资料来源：Wind，兴业证券经济与金融研究院整理

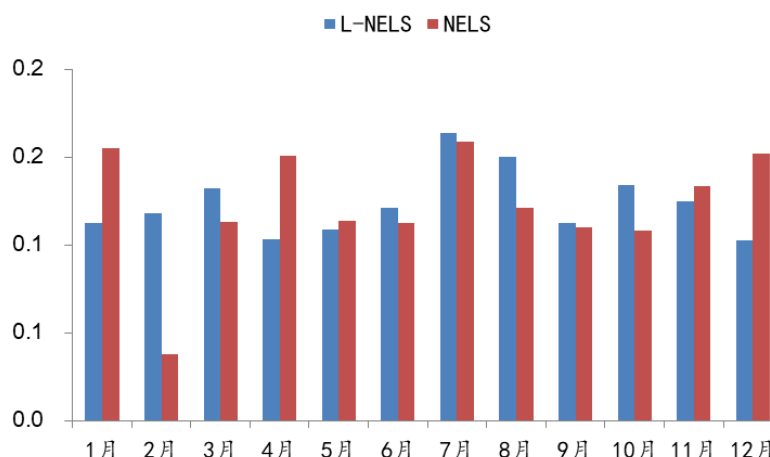
图表 30、L-NELS 因子多空净值曲线



资料来源：Wind，兴业证券经济与金融研究院整理

仅仅从数字来看，NELS 与 L-NELS 因子表现相差无几，但如果分月度来观察，就会发现两者有较大的差异。尤其是在 2 月份，基于 L-NELS 因子的 IC 均值达到 0.12，而 NELS 因子的 IC 仅仅为 0.04。其实我们都知道 A 股长期以来都呈现出较为明显的春节效应，如果不做特殊的处理，很多模型或者因子在 1、2 月份的表现都差强人意（基于 Adaboost 模型得到的 NELS 因子也不例外）。正是因为失败案例中 2 月份样本出现的次数较多，这才最终导致了 L-NELS 因子对春节效应（或 2 月效应）的把握更为精准。

图 31、NELS & L-NELS 因子在不同月份的 IC 表现



资料来源: Wind, 兴业证券经济与金融研究院整理

鉴于两者的互补性较强, 我们尝试了以下几种方式来利用 L-NELS 因子对 NELS 因子进行修正与增强:

- 1) 将 L-NELS 因子与 NELS 按照 1:2 合成, 称之为 NELS_Comb_Loss;
- 2) 将 NELS 因子在 1 月底的值替换成 L-NELS, 其他月份不变, 称之为 NELS_Loss_FebRep;
- 3) 将 NELS 在 1 月底的值替换成 L-NELS, 同时其他月份按照 1:2(L-NELS:NELS) 的比例合成, 称之为 E-NELS(Enhanced Neutralized Ensemble Learning Signal)。

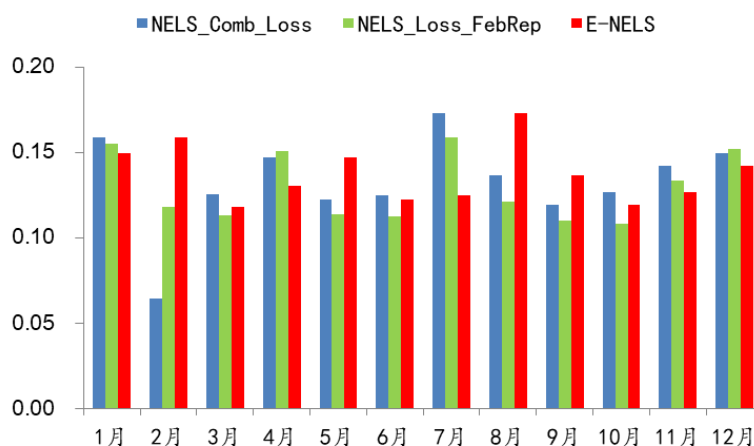
对比三个合成因子的表现: 第三种合成方式的因子 IC 达到 0.136、IC_IR 达到 2.02, 有效性要远高于方式 1 和方式 2 的表现。而分月度研究三个因子的 IC 表现, 与我们的预期也比较一致: 方式 3 合成因子在各个月度的表现更加平稳有效。

图 32、NELS 与 E-NELS 因子在不同情况下合成的 IC 检验

	平均值	Std	最小值	最大值	IC_IR	T 值
NELS_Comb_Loss	0.132	0.08	-0.15	0.34	1.76	15.1
NELS_Loss_FebRep	0.129	0.07	-0.10	0.33	1.74	14.8
E-NELS	0.136	0.07	-0.03	0.34	2.02	17.4

资料来源: Wind, 兴业证券经济与金融研究院整理

图 33、NELS 与 E-NELS 因子在不同情况下合成的 IC 月度均值



资料来源: Wind, 兴业证券经济与金融研究院整理

请务必阅读正文之后的信息披露和重要声明

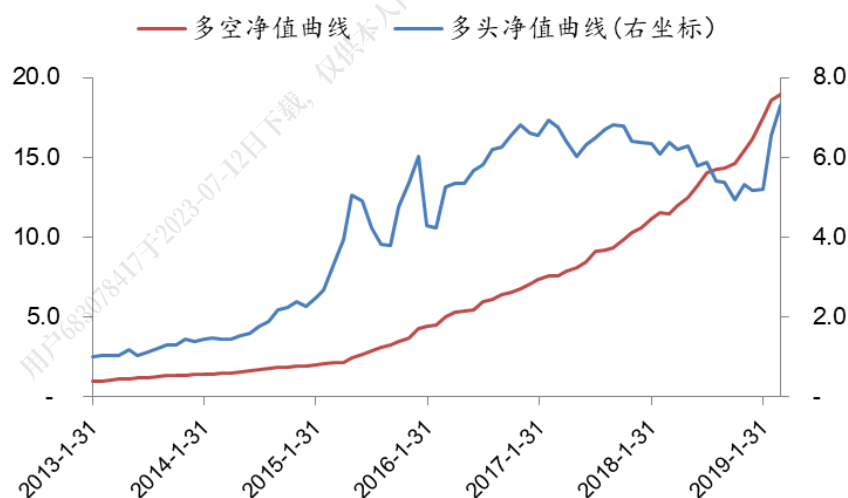
我们进一步研究 E-NELS 因子分位数组组合的表现。从测试结果来看，各组别严格单调，规律更加明显，多空组合的年化收益率高达 61.2%，夏普率达到 6.12，最大回撤仅为 0.5%，改进效果非常明显。因此在后续的研究中，如果没有特别说明，我们都将基于 E-NELS 因子展开。同时需要注意的是：在后续的研究中，我们将样本填充至 2006 年 1 月底，其中 2006-2012 年底这段数据将用原始的 NELS 因子来替代。

图表 34、E-NELS 分位数组组合测试结果

	总收益率	年化收益率	Sharpe 比率	换手率	最大回撤率	年化超额收益率	跟踪误差	信息比率	胜率
top	631.2%	38.1%	1.15	137%	29.7%	20.4%	5.3%	3.87	88.0%
1	401.2%	29.9%	0.93	166%	34.1%	13.1%	3.0%	4.35	90.7%
2	327.9%	26.6%	0.84	172%	34.4%	10.0%	3.3%	3.02	82.7%
3	298.8%	25.2%	0.79	174%	35.5%	8.8%	3.3%	2.68	82.7%
4	199.0%	19.5%	0.63	175%	44.8%	3.5%	2.9%	1.23	62.7%
5	167.7%	17.3%	0.56	175%	48.8%	1.5%	2.9%	0.52	56.0%
6	112.8%	13.0%	0.41	174%	57.4%	-1.8%	3.4%	-0.54	34.7%
7	41.8%	5.8%	0.19	172%	64.8%	-8.3%	2.9%	-2.85	16.0%
8	2.6%	0.4%	0.01	166%	71.5%	-13.1%	4.5%	-2.89	18.7%
bottom	-64.0%	-15.3%	-0.48	143%	87.1%	-26.6%	6.4%	-4.15	2.7%
市场	139.2%	15.2%	0.48		52.2%				
L_S	1796.4%	61.2%	6.12		0.5%				

资料来源：Wind，兴业证券经济与金融研究院整理

图表 35、E-NELS 因子多空净值曲线



资料来源：Wind，兴业证券经济与金融研究院整理

至此我们完成了对基本模型的增强与提升。开篇伊始我们曾提到，4 年前兴业金融工程团队就曾经利用传统 Adaboost 算法构建过类似的选股因子。对比本文构建的 E-NELS 因子与原有因子的表现，E-NELS 不论是在有效性还是稳定性方面，都要远优于原来的因子，这也从一个角度印证了我们增量研究工作的意义。

图表 36、E-NELS 因子与原有因子对比之 IC 分析

股票池	平均值	Std	最小值	最大值	IC_IR	T 值
原有因子	0.112	0.08	-0.22	0.31	1.47	18.5
E-ENLS	0.131	0.07	-0.03	0.34	1.91	24.1

注：测试时间窗口：2006 年 1 月-2019 年 4 月 30 日

资料来源：Wind，兴业证券经济与金融研究院整理

图表 37、E-NELS 因子与原有因子对比之分位数测试分析

	E-ENLS				原有因子			
	年化收益	Sharpe	换手率	回撤	年化收益	Sharpe	换手率	回撤
top	46.1%	1.26	144.0%	61.0%	39.9%	1.13	127.4%	61.8%
1	41.2%	1.13	168.7%	62.9%	35.0%	0.98	162.8%	61.9%
2	38.5%	1.06	173.7%	60.6%	31.0%	0.88	170.2%	64.6%
3	32.5%	0.89	176.5%	65.8%	26.4%	0.74	172.2%	66.5%
4	27.3%	0.76	176.9%	66.1%	27.3%	0.76	173.5%	66.5%
5	22.5%	0.63	176.5%	67.8%	29.3%	0.79	173.4%	66.5%
6	18.2%	0.51	175.9%	67.0%	19.0%	0.53	172.1%	66.2%
7	11.4%	0.32	173.9%	71.3%	13.7%	0.38	169.1%	68.8%
8	3.8%	0.11	168.0%	73.2%	7.6%	0.20	165.1%	72.6%
bottom	-9.9%	-0.27	147.0%	94.0%	-6.2%	-0.16	139.8%	89.9%
L-S	60.3%	6.41	4.1%		46.4%	4.63		9.0%

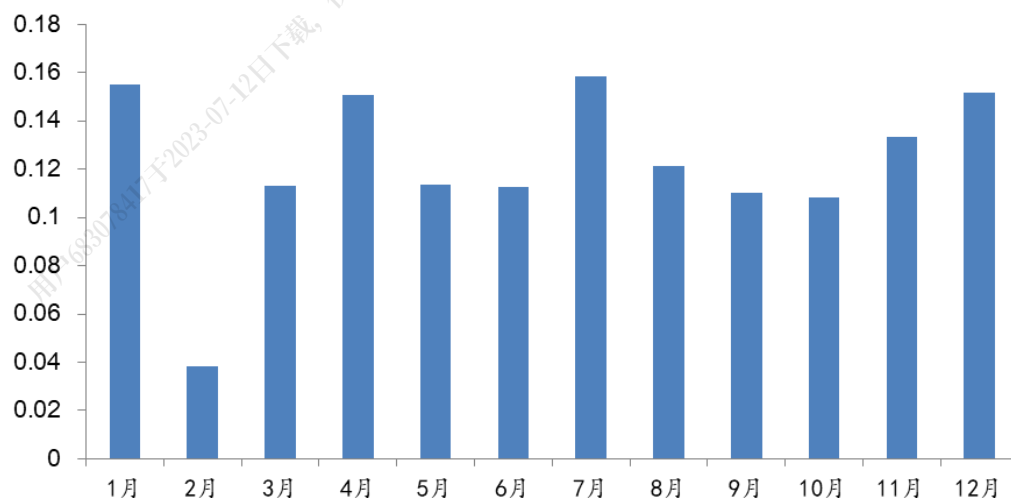
注：测试时间窗口：2006 年 1 月-2019 年 4 月 30 日

资料来源：Wind，兴业证券经济与金融研究院整理

3.2、月度效应

日历效应（calendar effect）是指金融市场与日期相联系的非正常收益，主要包括季节效应、月份效应、星期效应和假日效应，它们分别指金融市场与季节、月份、星期和假日有关的非正常收益。日历效应是一种通过有效市场理论无法做出合理解释的市场异象，海外对它的研究最早可以追溯到上世纪 30 年代。前面展示了 NELS 因子的月度 IC 表现，从结果来看，不同月份表现差异较大，尤其是 2 月份，模型在该月份的选股能力显著落后于其他月份。

图表 38、NELS 因子月度 IC 表现



资料来源：Wind，兴业证券经济与金融研究院整理

为了捕捉不同月份的风格特征，我们考虑使用多年同一月份的数据来训练样本。具体来讲，在每个月末，我们使用过去一段时间窗口中与当前研究时点相同月份的数据作为算法的训练集，其他设置不变。这里我们使用过去 10 年的相同月份数据来训练模型，同时要求过去 10 个相同月份中至少 8 个月有数据。因此，该因子从 2013 年 1 月份开始有值。我们称此模型输出的复合因子为 NELS_Mth 因子，进一步研究该因子的选股能力。

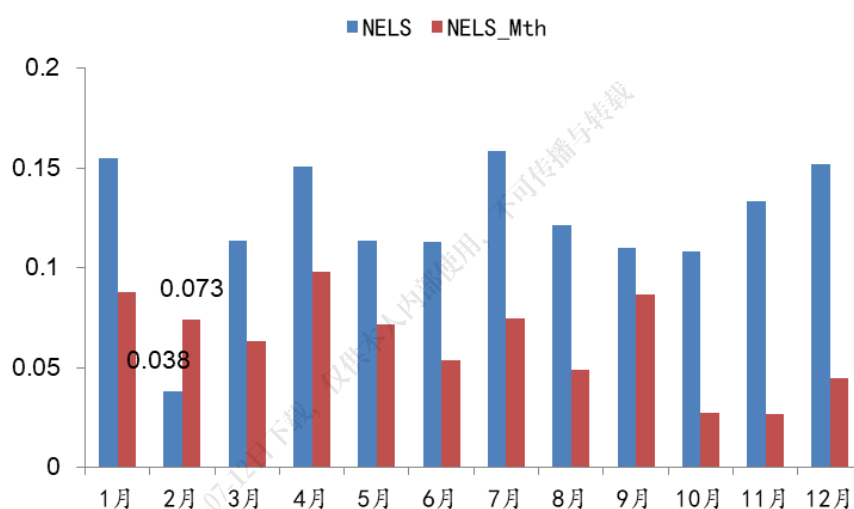
对比同期的 NELS 因子以及 NELS_Mth 因子的表现，我们发现虽然 NELS_Mth 因子的综合选股能力不如 NELS 因子，但是在 2 月份上确有显著的改善（参见图表 39、40）：IC 直接从 0.038 提升至 0.073。因此我们将两个因子合成：将 NELS 在 1 月份的值替换成 NELS_Mth，其他月份保持不变，将新合成的因子称之为 NELS_Mth_Rep。

图表 39、NELS 以及 NELS_Mth 因子 IC 检验

	平均值	Std	最小值	最大值	IC_IR	T 值
NELS	0.122	0.086	-0.26	0.33	1.40	12.1
NELS_Mth	0.063	0.087	-0.18	0.22	0.72	6.17

资料来源：Wind，兴业证券经济与金融研究院整理

图表 40、NELS 以及 NELS_Mth 因子的月度 IC 表现



资料来源：Wind，兴业证券经济与金融研究院整理

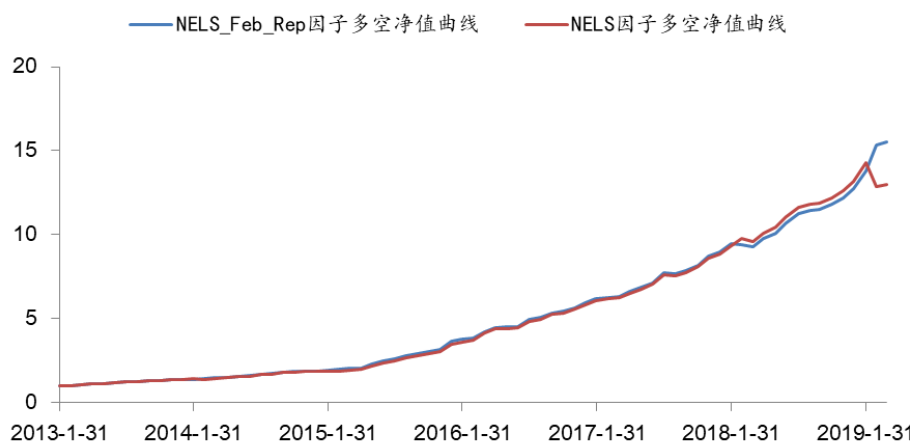
对比 NELS_Feb_Rep 与 NELS 因子自 2013 年 1 月份以来的表现，我们发现经过改进的因子表现更加稳定，也有效的规避了 NELS 在 2019 年 2 月份的大幅回撤。

图表 41、NELS 以及 NELS_Feb_Rep 因子 IC 对比

	平均值	Std	最小值	最大值	IC_IR	T 值
NELS	0.122	0.086	-0.26	0.33	1.40	12.1
NELS_Feb_Rep	0.125	0.081	-0.10	0.33	1.54	13.12

资料来源：Wind，兴业证券经济与金融研究院整理

图表 42、NELS 以及 NELS_Feb_Rep 因子的多空净值曲线



资料来源：Wind，兴业证券经济与金融研究院整理

3.3、本章小结

在第二章中我们详细介绍了如何利用改进版的Adaboost算法构建量化选股模型，并进一步研究了行业市值中性化对因子的影响、模型的优势、不同时间点算法所选风格因子的差异等议题；而来到第三章，我们把关注点放在了如何利用失败者样本和月度效应对模型进行改进和增强上，应该说最终的结果是令人满意的。不过需要指出的是，虽然基于日历效应（3.2 章节）的改进方法对原有的 NELS 因子也有一定的提升，但其表现不如基于失败者样本的改进效果。为了方便和简化研究，后续我们将只关注 E-NELS 因子的相关结果。

4、高低换手风格选股模型搭建

月度，应该说是当前市场最为流行的量化股票类策略的调仓频率，其在信息利用的及时性和组合管理的便利性之间找到了一个较好的平衡点。但组合管理者在实践中遇到的问题总是有其各自的考量。对于组合规模较大或公司对交易及风险把控较严的管理人来说，月度调仓有时还是太过频繁，特别是当市场流动性一般时，往往会给模型的实现带来比较大的压力；而对于一些管理资金规模并不是太大，或其本身来自于非公募保险机构，交易制度较为灵活的管理人来说，提高换手频率，更加充分地去把握一些中短期投资机会以提高组合的收益率，往往是他们的核心诉求。有鉴于此，我们在前文所开发的月度量化模型的基础上，又搭建了两类换手风格差异明显的选股模型，以供投资者参考。

4.1、低换手模型构建

前面我们已经阐述了对低换手模型需求的背景，接下来就尝试通过不同的方法降低模型换手率。我们的主要方法有以下两种：

- 1、优化所选因子：剔除因子库中的高换手率因子，保留低换手因子来进行模型的训练和学习，以达到降低换手率的目的；
- 2、延长向前预测期：将对未来 1 个月进行预测，调整为对未来 3 个月进行预

测，以使得更多具有长期预测能力的因子进入模型，从而降低换手率。

1. 优化所选因子

在前述研究中我们提到六大类因子共计 165 个（价值 19 个、成长 33 个、动量反转 9 个、情绪 13 个、另类 52 个、质量 39 个）。因子逻辑不同，其换手率也有很大的差异，比如价值类因子的换手率相对较低，一般多头单边换手率不超过 30%，而另类和动量反转因子换手率相对就高一些。我们用因子前后两期的 spearman 相关性作为换手率衡量的指标（一般高换手率的因子前后期的相关性必然很低），研究 165 个因子自相关性大小，将自相关性小于阈值的因子删除，用余下的指标作为模型的输入构建低换手选股模型。

图表 43、不同风格因子的自相关性系数均值

类别	成长	动量反转	价值	另类	情绪	质量
自相关系数	77.5%	55.0%	89.4%	58.0%	41.9%	83.4%

资料来源：Wind，兴业证券经济与金融研究院整理

我们在因子自相关性阈值分别设定为 75%、80%、85%、90% 的情形下，均计算了相应的 E-NELS 因子的表现。

图表 44、不同阈值下的因子池总数

阈值	成长	动量反转	价值	另类	情绪	质量	因子总数
75%	17	5	16	21	2	32	93
80%	14	4	16	20	2	26	82
85%	12	3	16	20	1	17	69
90%	7	2	15	18	1	10	53

资料来源：Wind，兴业证券经济与金融研究院整理

当阈值自 75% 上升至 85% 时，因子的选股能力并没有显著变化，同时自相关性水平以及换手率也得到了有效的控制，而当阈值进一步提升至 90% 时，虽然换手率和自相关性得到了更为明显的改善，但是选股能力也大幅度下降。综合考虑之后，我们将阈值设定在了 80%。

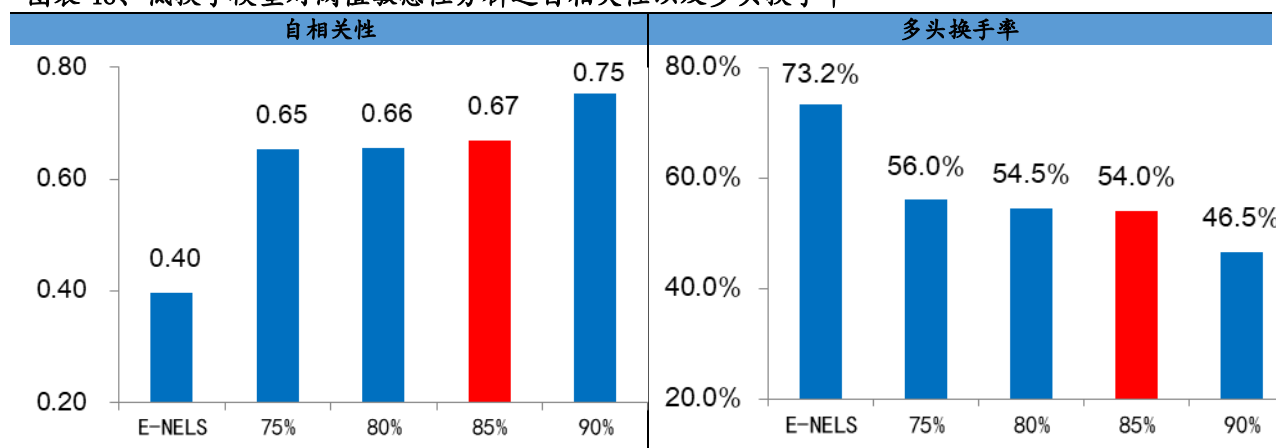
此时，通过删除高换手率因子所得到 LowTurn_E-NELS 因子的自相关性由 0.4 提升至 0.67，多头组合的单边换手率自 73.2% 降低至 54.0%，边际改善非常显著。

图表 45、低换手模型对阈值敏感性分析之 IC 检验

阈值	平均值	Std	最小值	最大值	IC_IR	t 统计量
E-NELS	0.136	0.07	-0.03	0.34	2.02	17.4
75%	0.104	0.07	-0.09	0.26	1.48	18.57
80%	0.099	0.07	-0.11	0.27	1.35	16.94
85%	0.096	0.08	-0.10	0.28	1.23	15.52
90%	0.079	0.08	-0.15	0.27	1.01	12.75

资料来源：Wind，兴业证券经济与金融研究院整理

图表 46、低换手模型对阈值敏感性分析之自相关性以及多头换手率



资料来源：Wind，兴业证券经济与金融研究院整理

2. 延长向前预测期

在机器学习过程之中，如果预测的标的特征有变动，那么最终选择的因子必然也是能反应这种变化的。假设我们让预测窗口从 1 个月延长至季度（3 个月），那么模型会自适应的去选择一些 IC 衰减相对较慢的因子，自然而然地策略换手率也会随之降低。基于这个思想，我们也训练了相应的模型。

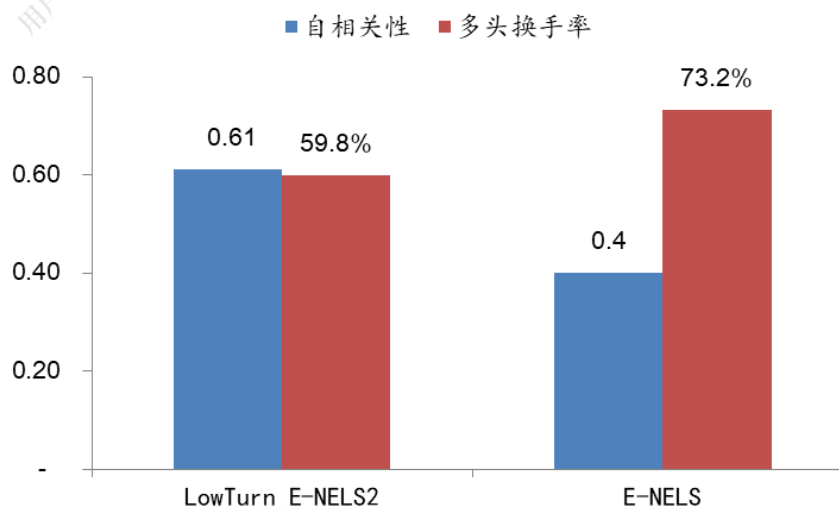
从测试结果来看，LowTurn_E-NELS2 因子选股有效性也非常强，同时因子自相关性为 0.62，多头组合的单边换手率降低至 60%，相对于原始的 E-NELS 因子均有一定的改善和提升。

图表 47、预测窗口延长下的低换手因子 IC 表现

	平均值	Std	最小值	最大值	IC_IR	t 统计量
E-NELS	0.136	0.07	-0.03	0.34	2.02	17.4
LowTurn_E-NELS2	0.102	0.09	-0.14	0.32	1.16	14.5

资料来源：Wind，兴业证券经济与金融研究院整理

图表 48、预测窗口延长下的低换手因子自相关性以及多头换手率



资料来源：Wind，兴业证券经济与金融研究院整理

从测试结果来看，两种改进方式都达到了预期的目的和效果：第一种方法换手率控制的更稳定；第二种通过机器学习的方式进行自适应调整，并没有完全控制住换手，但仍然不失为一种有效的尝试方法。

4.2、高换手模型构建

市场云谲波诡，因子的有效性时常稍纵即逝，很多人都希望尽量多的去把握有限的投资机会，斩获投资收益。传统量化选股模型往往以月度或季度频率调仓，这种做法有利于大资金的管理，但随着量化投资的普及和市场竞争的加剧，低频组合管理的劣势也越来越多的体现了出来：因子规律外推不稳定、模型同质化导致的策略拥挤愈加普遍...有鉴于此，我们试图通过提高模型换手频率，以表现更为稳定的中短期价量因子为基础，打造一套周度模型来应对目前的某些困难。在本节中，仿照前几章中月度机器学习模型的搭建方式，我们将使用过去 52 周的数据来预测未来 1 周股票收益的表现。

这里面需要注意的是，在周度调仓的情况下，如价值、成长、质量等数据本身更新频率较低的因子类别可能就不太适用了。因此在构建周度选股模型时，我们仅考虑了动量反转类以及另类因子。

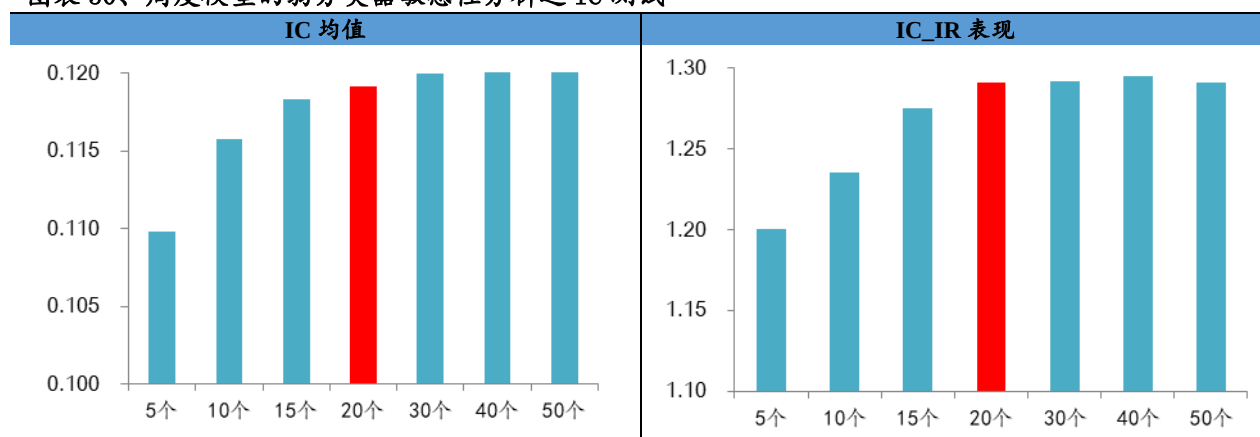
同前面的月度 NELS 因子研究一致，对于周度 NELS 因子，我们遍历了弱分类器为 5、10、15……60 情况下因子的表现。从 IC 测试结果来看，当弱分类器数目从 5 提升至 20 时，IC、IC_IR 边际改善很明显，而后提升效果有限；从分位数测试结果来看，也有着相同的规律。所以，后续对周度因子的测试都将基于弱分类器数量为 20 的结果来进行。

图表 49、周度选股因子的 IC 敏感性分析

	平均值	Std	最小值	最大值	IC_IR	t 统计量
NELS_Week_5	0.110	0.09	-0.22	0.43	1.20	31.0
NELS_Week_10	0.116	0.09	-0.19	0.48	1.24	32.0
NELS_Week_15	0.118	0.09	-0.20	0.48	1.28	33.0
NELS_Week_20	0.119	0.09	-0.22	0.48	1.29	33.4
NELS_Week_30	0.120	0.09	-0.22	0.47	1.29	33.4
NELS_Week_40	0.120	0.09	-0.22	0.48	1.29	33.5
NELS_Week_50	0.120	0.09	-0.22	0.47	1.29	33.4

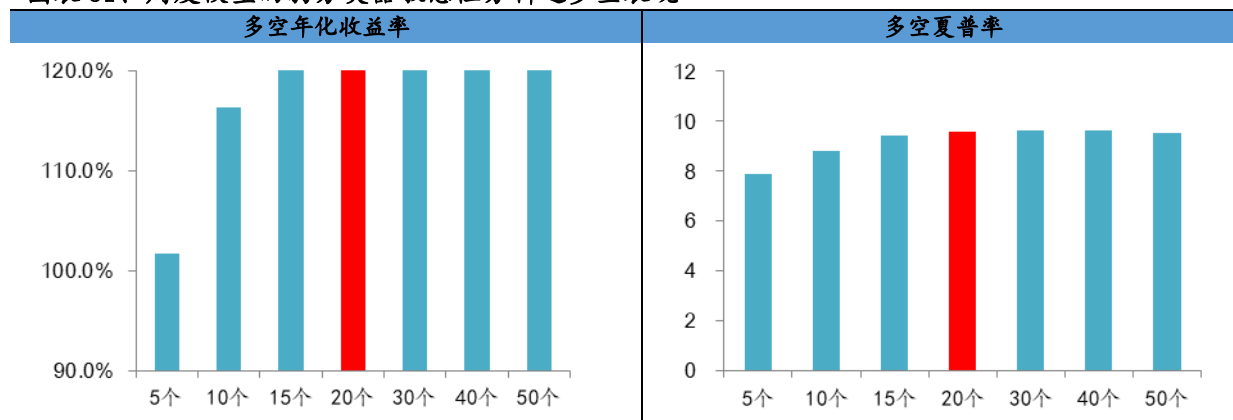
资料来源：Wind，兴业证券经济与金融研究院整理

图表 50、周度模型的弱分类器敏感性分析之 IC 测试



资料来源：Wind，兴业证券经济与金融研究院整理

图表 51、周度模型的弱分类器敏感性分析之多空表现



资料来源：Wind，兴业证券经济与金融研究院整理

不难发现 E-NELS_Week 与 NELS_Week 表现的整体差异较小，但 E-NELS_Week 多空组合的回撤要远小于 NELS_Week 因子。因此在后文中，我们将主要讨论和使用 E-NELS_Week 因子。

图表 52、周度因子的 IC 检验

	平均值	Std	最小值	最大值	IC_IR	t 统计量
NELS_Week	0.119	0.09	-0.22	0.47	1.30	33.76
E-NELS_Week	0.114	0.09	-0.21	0.42	1.30	33.96

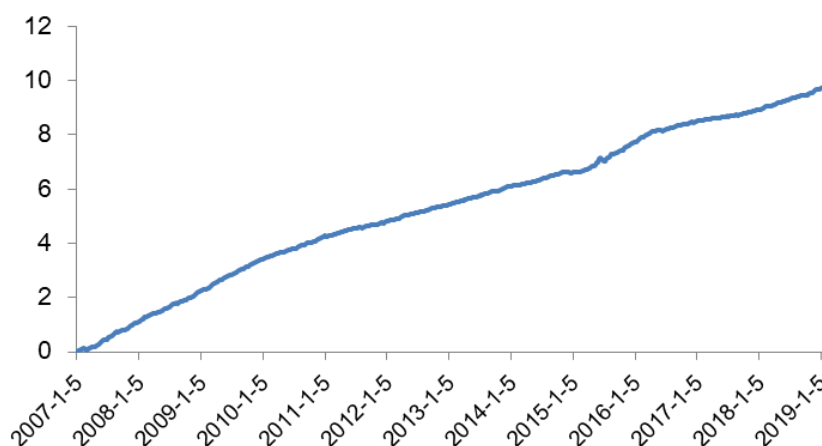
资料来源：Wind，兴业证券经济与金融研究院整理

图表 53、NELS_Week Vs. E-NELS_Week 因子分位数组合测试结果

	NELS_Week 因子表现				E-NELS_Week 因子表现			
	年化收益	Sharpe	换手率	回撤	年化收益	Sharpe	换手率	回撤
top	69.0%	1.96	141.8%	58.7%	68.3%	2.04	138.6%	57.7%
1	52.7%	1.52	167.2%	60.9%	50.9%	1.56	165.3%	61.3%
2	44.5%	1.31	172.8%	63.9%	44.3%	1.39	171.4%	63.1%
3	38.7%	1.17	174.5%	63.5%	39.8%	1.27	173.6%	63.5%
4	28.6%	0.89	174.3%	66.6%	31.7%	1.01	173.6%	64.9%
5	20.7%	0.66	173.8%	69.7%	23.8%	0.76	173.9%	67.9%
6	14.5%	0.47	173.4%	68.3%	16.2%	0.51	173.6%	70.0%
7	7.5%	0.25	172.1%	73.1%	4.1%	0.13	172.3%	75.6%
8	-4.7%	-0.15	168.1%	88.1%	-5.9%	-0.18	168.2%	90.8%
bottom	-25.3%	-0.78	148.3%	99.3%	-26.3%	-0.76	147.4%	99.4%
市场	22.1%	0.69		68.6%	22.1%	0.69		68.6%
L-S	125.2%	9.59		15.4%	124.5%	9.82		6.4%

资料来源：Wind，兴业证券经济与金融研究院整理

图表 54、E-NELS_Week 因子的多空净值曲线（对数化）



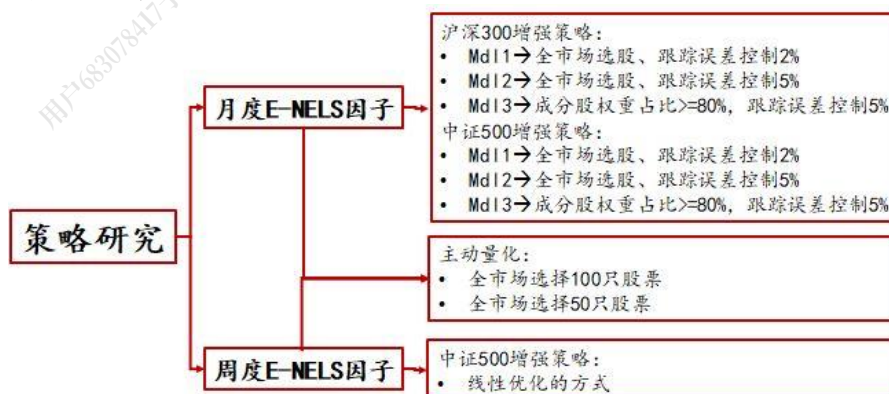
资料来源：Wind，兴业证券经济与金融研究院整理

5、股票模拟组合体系搭建

抽丝剥茧，我们一步步论证了改进版 Adaboost 算法在量化选股层面的有效性；万剑归宗，再好的研究也要经受模拟组合乃至实盘交易的考验，不然终归不过是纸上谈兵。

一般而言，投资者会结合自己的需求（如投资风险偏好不同、资金容量要求不同、换手频率不同等）构建不同风格的投资策略。有鉴于此，我们亦构建了不同的模拟组合，供大家参考。具体框架参见图表 55，策略具体描述请参考后文相应章节。本章中的测试时间窗口统一为：2009 年 12 月 31 日-2019 年 5 月 31 日。

图表 55、策略层面的研究框架



资料来源：兴业证券经济与金融研究院整理

5.1、因子在沪深 300&中证 500 股票池内选股表现

在针对沪深 300 以及中证 500 做指数增强以及相关策略研究之前，我们首先测试月度、周度 E-NELS 因子在两个宽基指数范围内的选股有效性。从因子层面的测试效果来看，无论是周度还是月度，E-NELS 因子在沪深 300 以及中证 500 成分股内部的选股有效性是非常强的：以月度 E-NELS 因子在沪深 300 股票池内

部表现为例（其他测试结果请见图表 56、57、58），因子 IC 达到 0.09，T 值显著性高达 12.2，分位数组合严格单调，多空组合年化收益率高达 38.4%，夏普率达到 2.83。

E-NELS 在全市场以及沪深 300、中证 500 各个股票池测试的有效性说明了机器学习模型的稳定性，为进一步开发基于这些指数的增强选股策略打下了坚实的基础。

图表 56、周度&月度 E-NELS 因子在沪深 300 及中证 500 股票池内 IC 测试

	股票池	平均值	Std	最小值	最大值	IC_IR	T 值
月度 E-NELS	沪深 300	0.091	0.09	-0.10	0.28	0.99	12.2
月度 E-NELS	中证 500	0.127	0.08	-0.03	0.37	1.60	19.0
周度 E-NELS	沪深 300	0.072	0.10	-0.28	0.35	0.72	18.0
周度 E-NELS	中证 500	0.107	0.10	-0.23	0.41	1.09	27.1

资料来源：Wind，兴业证券经济与金融研究院整理

图表 57、月度 E-NELS 因子在沪深 300 及中证 500 股票池内分位数测试

沪深 300 内测试效果					中证 500 内测试效果			
	年化收益	Sharpe	换手率	回撤	年化收益	Sharpe	换手率	回撤
top	28.6%	0.84	156.3%	59.4%	28.9%	0.83	144.9%	60.7%
1	23.1%	0.68	170.9%	63.3%	22.5%	0.64	159.2%	63.1%
2	20.1%	0.61	167.5%	68.7%	20.5%	0.60	158.7%	64.6%
3	19.3%	0.59	164.9%	66.8%	18.4%	0.52	158.6%	65.6%
4	15.6%	0.47	160.6%	69.1%	13.8%	0.40	158.1%	68.2%
5	12.9%	0.39	162.1%	69.4%	9.7%	0.28	159.4%	70.0%
6	9.8%	0.29	167.9%	69.5%	5.4%	0.16	159.2%	69.1%
7	5.9%	0.18	168.4%	70.3%	0.3%	0.01	161.8%	74.6%
8	-3.2%	-0.10	170.3%	85.5%	-6.8%	-0.20	158.0%	84.0%
bottom	-8.7%	-0.25	159.2%	93.3%	-10.3%	-0.30	146.3%	89.8%
市场	13.1%	0.40		69.0%	10.4%	0.30		68.4%
L-S	38.4%	2.83		7.1%	42.3%	3.82		4.2%

资料来源：Wind，兴业证券经济与金融研究院整理

图表 58、周度 E-NELS 因子在沪深 300 及中证 500 股票池内分位数测试

沪深 300 内测试效果					中证 500 内测试效果			
	年化收益	Sharpe	换手率	回撤	年化收益	Sharpe	换手率	回撤
top	27.3%	0.91	140.1%	59.9%	48.1%	1.45	136.1%	60.9%
1	18.5%	0.62	166.4%	63.2%	31.6%	0.94	165.3%	62.1%
2	15.3%	0.52	172.6%	66.1%	26.7%	0.82	172.0%	64.5%
3	12.1%	0.41	175.8%	65.7%	23.0%	0.70	174.6%	65.2%
4	10.9%	0.37	176.6%	70.0%	18.2%	0.56	175.5%	68.0%
5	6.2%	0.22	176.1%	74.1%	10.1%	0.31	175.5%	68.7%
6	1.7%	0.06	174.9%	72.0%	4.5%	0.14	174.6%	73.3%
7	-0.1%	-0.00	172.6%	71.4%	-1.6%	-0.05	173.0%	74.8%
8	-10.5%	-0.35	167.8%	86.3%	-10.4%	-0.31	168.0%	87.7%
bottom	-18.3%	-0.58	141.5%	96.1%	-28.3%	-0.81	143.5%	99.1%
市场	5.7%	0.20		70.3%	10.4%	0.32		69.4%
L-S	53.2%	3.89		15.7%	102.2%	7.49		6.8%

资料来源：Wind，兴业证券经济与金融研究院整理

5.2、基于月度 E-NELS 因子的增强策略研究

指数增强策略旨在控制跟踪误差的前提下，尽可能获取超越基准的表现。综合考虑投资者的偏好，针对于一个宽基指数（沪深 300/中证 500），我们构建了 3 个增强模型，不同模型的实现细节参见图表 59。

图表 59、基于月度 E-NELS 因子构建的增强策略框架

	针对于沪深 300 的增强策略	针对于中证 500 的增强策略
策略一	1、跟踪误差不超过 2%； 2、换手率约束单边 0.2	1、跟踪误差不超过 2%； 2、换手率约束单边 0.4
策略二	1、跟踪误差不超过 5%； 2、换手率约束单边 0.2	1、跟踪误差不超过 5%； 2、换手率约束单边 0.4
策略三	1、跟踪误差不超过 5%； 2、换手率约束单边 0.2 3、每期沪深 300 股票池内股票权重占 比>80%	1、跟踪误差不超过 5%； 2、换手率约束单边 0.4 3、每期中证 500 股票池内股票权重 占比>80%
统一要求	1、选股池需删除 ST 股、同时保证上市天数满 180 天； 2、保持行业市值中性； 3、个股最大权重不超过 10%； 4、交易成本：0.3%	

资料来源：兴业证券经济与金融研究院整理

➤ 基于月度 E-NELS 因子的沪深 300 增强策略

基于沪深 300 的三个增强策略表现非常优异，策略相对沪深 300 指数的年化超额收益稳定在 7%-8%，最大回撤徘徊在 2.5%-2.8%，风险收益比高达 3.3。分年度来看，选股策略在每一年均能稳定的跑赢基准，2017-2018 年间年化超额收益稳定在 5%左右，2019 年以来超额收益稳定在 2%左右。

图表 60、针对于沪深 300 的增强策略表现汇总

	年化收益率	波动率	风险收益比	胜率	回撤
沪深 300 表现	0.2%	23.1%	0.01	51.1%	46.7%
策略一多头	8.0%	22.7%	0.35	52.4%	41.1%
策略二多头	8.3%	22.8%	0.37	52.6%	40.4%
策略三多头	8.5%	22.8%	0.37	52.5%	39.8%
策略一超额	7.7%	2.3%	3.34	59.8%	2.5%
策略二超额	8.0%	2.5%	3.17	59.0%	2.7%
策略三超额	8.2%	2.5%	3.31	59.8%	2.5%

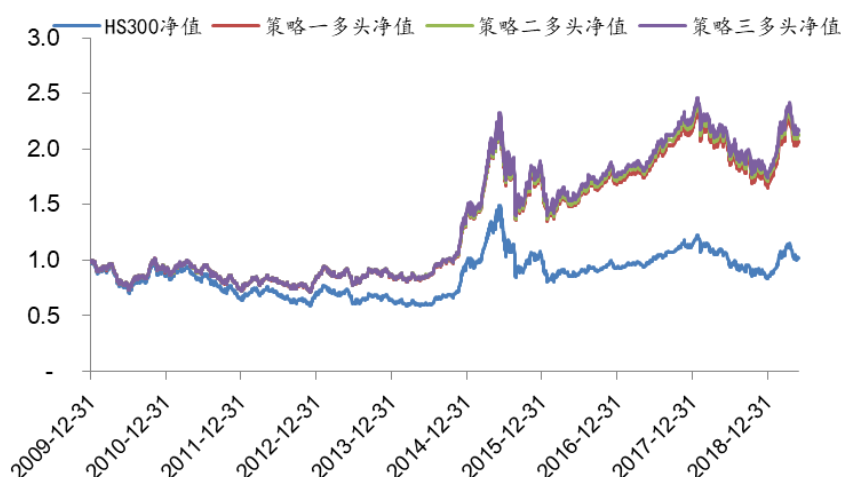
资料来源：Wind，兴业证券经济与金融研究院整理

图表 61、针对于沪深 300 构建的增强策略分年度表现

	基准	多头			超额		
		策略一	策略二	策略三	策略一	策略二	策略三
2010 年	-12.5%	-8.0%	-7.7%	-7.7%	4.7%	5.0%	5.1%
2011 年	-25.0%	-19.1%	-19.2%	-19.2%	7.8%	7.7%	7.7%
2012 年	7.6%	14.7%	14.7%	14.6%	6.5%	6.5%	6.4%
2013 年	-7.6%	2.2%	2.4%	2.5%	10.4%	10.6%	10.7%
2014 年	51.7%	63.5%	64.3%	66.2%	7.7%	8.2%	9.6%
2015 年	5.6%	23.5%	24.4%	25.8%	17.1%	17.9%	19.0%
2016 年	-11.3%	-4.1%	-3.8%	-4.1%	8.2%	8.6%	8.2%
2017 年	21.8%	27.8%	28.5%	28.6%	5.0%	5.6%	5.8%
2018 年	-25.3%	-22.4%	-21.7%	-21.8%	3.8%	4.7%	4.5%
2019 年	20.6%	23.2%	22.3%	22.4%	2.1%	1.3%	1.4%

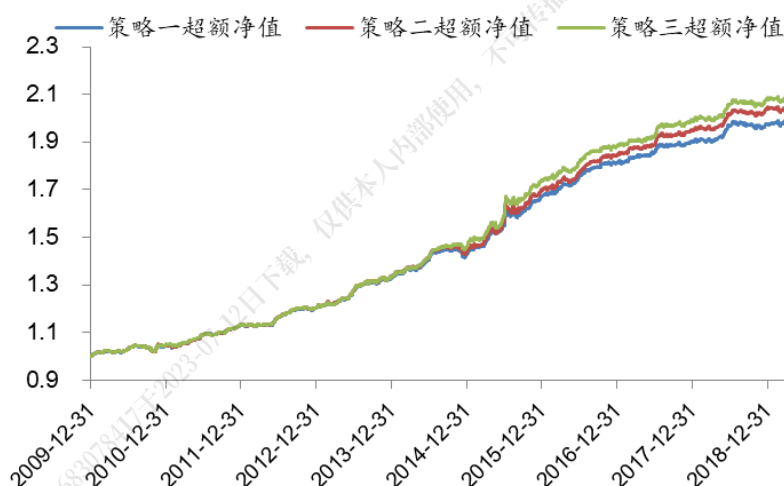
资料来源：Wind，兴业证券经济与金融研究院整理

图表 62、针对于沪深 300 构建的增强策略多头净值曲线



资料来源：Wind，兴业证券经济与金融研究院整理

图表 63、针对于沪深 300 构建的增强策略超额净值曲线



资料来源：Wind，兴业证券经济与金融研究院整理

➤ 基于月度 E-NELS 因子的中证 500 增强策略

与沪深 300 的增强策略框架一致，我们构建了中证 500 指数增强策略。从结果来看，策略表现突出。以策略二为例：策略年化超额收益率高达 16.2%，风险收益比达到 4.08，回撤仅仅为 3.1%。分年度来看，各策略也能稳定的战胜基准，在 2017-2019 年期间分别能跑赢基准 5 个点左右。

图表 64、针对于中证 500 的增强策略表现汇总

	年化收益率	波动率	风险收益比	胜率	回撤
中证 500 表现	1.0%	26.8%	0.04	54.2%	65.2%
策略一多头	13.6%	26.8%	0.51	55.2%	50.6%
策略二多头	17.4%	26.3%	0.66	55.7%	46.2%
策略三多头	14.0%	26.5%	0.53	55.5%	46.8%
策略一超额	12.6%	3.0%	4.27	62.0%	2.1%
策略二超额	16.2%	4.0%	4.08	59.9%	3.1%
策略三超额	12.7%	3.5%	3.59	59.6%	2.7%

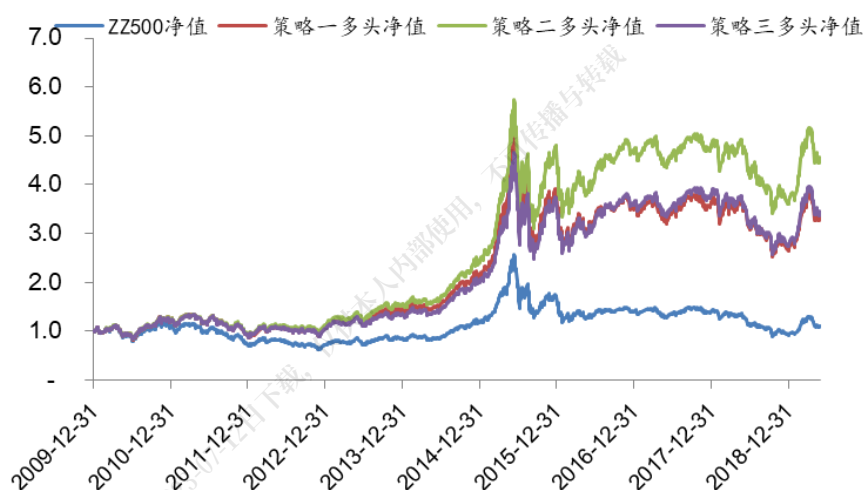
资料来源：Wind，兴业证券经济与金融研究院整理

图表 65、针对于中证 500 构建的增强策略分年度表现

	中证 500	多头			超额		
		策略一	策略二	策略三	策略一	策略二	策略三
2010 年	10.1%	24.1%	24.8%	22.9%	12.8%	13.6%	11.5%
2011 年	-33.8%	-25.6%	-24.1%	-24.6%	12.0%	13.7%	13.0%
2012 年	0.3%	17.9%	20.5%	14.7%	17.2%	19.7%	13.9%
2013 年	16.9%	32.5%	36.8%	27.3%	13.4%	16.8%	8.7%
2014 年	39.0%	47.2%	55.1%	49.1%	6.0%	11.9%	7.4%
2015 年	43.1%	81.2%	94.9%	81.9%	28.2%	37.6%	27.4%
2016 年	-17.8%	-9.0%	-1.5%	-2.5%	11.4%	19.7%	18.6%
2017 年	-0.2%	2.9%	2.5%	5.8%	3.2%	2.5%	5.9%
2018 年	-33.3%	-25.8%	-23.7%	-27.3%	11.0%	14.1%	8.8%
2019 年	17.8%	24.4%	24.8%	24.8%	5.5%	5.8%	5.9%

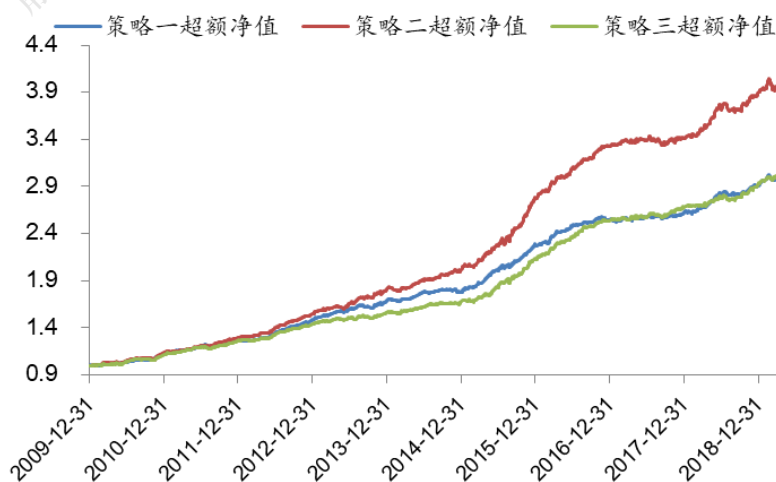
资料来源：Wind，兴业证券经济与金融研究院整理

图表 66、针对于中证 500 构建的增强策略多头净值曲线



资料来源：Wind，兴业证券经济与金融研究院整理

图表 67、针对于中证 500 构建的增强策略超额净值曲线



资料来源：Wind，兴业证券经济与金融研究院整理

5.3、基于月度 E-NELS 因子的主动量化策略研究

接下来，基于月度 E-NELS 因子，以收益最大化为目标，我们构建了两个主动量化选股模型，两个策略分别固定持仓 100 只和 50 只股票，策略实现细节参见图表 68。

图表 68、基于月度 E-NELS 因子构建的主动量化策略

	策略一	策略二
差异性要求	1、每期选择 100 只股票； 2、调仓日：若当期持仓的股票下一期没有跌出前 400，仍然继续持有	1、每期选择 50 只股票； 2、调仓日：若当期持有的股票下一期没有跌出前 200，仍然继续持有
统一要求	选股池需删除 ST、同时保证上市天数满 180 天；交易成本：0.3%	

资料来源：兴业证券经济与金融研究院整理

从测试结果来看，我们构建的主动量化策略成绩斐然，以选择 100 只股票的策略一为例，设定比较基准为中证 500：该策略年化超额收益率 25.3%，风险收益比高达 4.44。分年度来看，除 2017 年外，其他年份多头策略均能大幅跑赢基准。

图表 69、月度主动量化策略表现汇总

	年化收益率	波动率	风险收益比	胜率	回撤
中证 500 表现	1.0%	26.8%	0.04	54.2%	65.2%
策略一多头	26.1%	27.7%	0.94	56.7%	46.3%
策略二多头	27.0%	27.9%	0.97	57.1%	47.2%
策略一超额	25.3%	5.7%	4.44	61.2%	11.0%
策略二超额	26.1%	6.6%	3.93	60.2%	10.4%

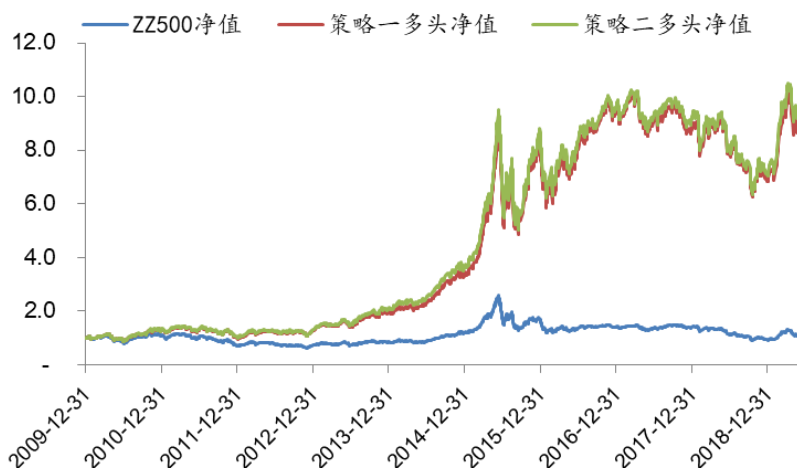
资料来源：Wind，兴业证券经济与金融研究院整理

图表 70、月度主动量化策略分年度表现

	中证 500	策略一多头	策略二多头	策略一超额	策略二超额
2010 年	10.1%	30.2%	32.2%	18.5%	20.1%
2011 年	-33.8%	-21.3%	-20.4%	17.8%	19.0%
2012 年	0.3%	26.5%	23.6%	25.8%	23.2%
2013 年	16.9%	51.5%	62.2%	30.3%	39.8%
2014 年	39.0%	69.8%	71.1%	23.0%	23.7%
2015 年	43.1%	150.1%	139.7%	78.7%	71.1%
2016 年	-17.8%	12.2%	10.4%	37.0%	33.8%
2017 年	-0.2%	-6.2%	-4.1%	-6.0%	-4.0%
2018 年	-33.3%	-21.7%	-22.7%	17.2%	15.9%
2019 年	17.8%	30.0%	34.4%	10.6%	14.3%

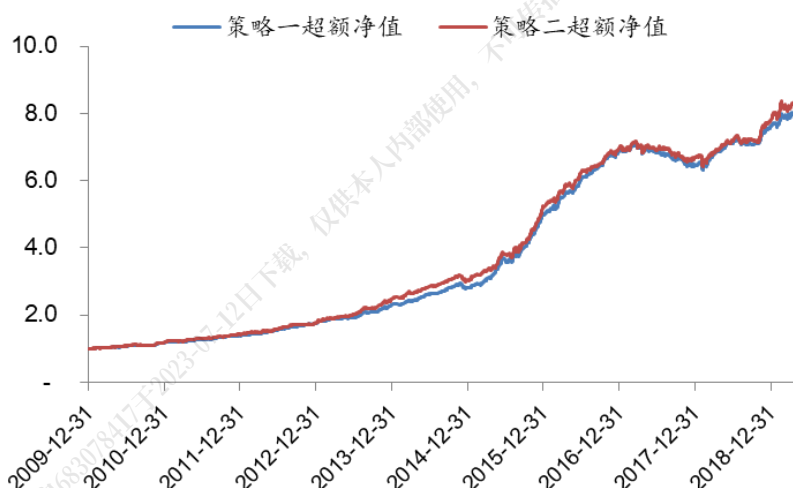
资料来源：Wind，兴业证券经济与金融研究院整理

图表 71、月度主动量化策略多头净值曲线



资料来源：Wind，兴业证券经济与金融研究院整理

图表 72、月度主动量化策略超额净值曲线



资料来源：Wind，兴业证券经济与金融研究院整理

5.4、基于周度 E-NELS 因子的增强模型研究

周度 E-NELS 因子主要运用的是一些价量指标训练生成，因而其对中小盘股票有效性更强，所以我们在利用周度 E-NELS 因子构建增强模型时主要考虑针对中证 500 的增强策略。策略实现细节如下：

- ✧ 目标：超额收益率最大化；
- ✧ 权重约束：单个股票的权重不超过 10%，同时相对基准的权重波动不超过 $\pm 0.5\%$ ；
- ✧ 换手率约束：周度单边 40 %；
- ✧ 其他约束：行业、市值、Beta 中性；交易成本 0.3%；

从表现来看，策略相对中证 500 指数的年化超额收益达到 21.6%，夏普率高达 4.26，最大回撤仅为 6%，策略表现非常优异。分年度来看，周度选股策略均能稳定的战胜基准，2019 年前 5 个月模型已跑赢中证 500 达 7.6%。

请务必阅读正文之后的信息披露和重要声明

图表 73、周度 E-NELS 因子增强策略整体表现

	总收益	年收益	波动率	夏普率	胜率	回撤
多头	611.1%	23.2%	25.3%	0.92	56.4%	42.1%
基准	9.5%	1.0%	26.8%	0.04	54.2%	65.2%
超额	530.2%	21.6%	5.1%	4.26	60.7%	6.0%

资料来源：Wind，兴业证券经济与金融研究院整理

图表 74、周度 E-NELS 因子增强策略净值曲线



资料来源：Wind，兴业证券经济与金融研究院整理

图表 75、周度 E-NELS 因子增强策略分年度表现

	基准	多头	超额
2010 年	10.1%	47.5%	34.1%
2011 年	-33.8%	-20.5%	19.6%
2012 年	0.3%	19.3%	19.0%
2013 年	16.9%	43.0%	22.0%
2014 年	39.0%	59.3%	14.6%
2015 年	43.1%	111.6%	46.2%
2016 年	-17.8%	2.1%	23.2%
2017 年	-0.2%	2.6%	2.5%
2018 年	-33.3%	-20.7%	18.8%
2019 年	17.8%	27.1%	7.6%

资料来源：Wind，兴业证券经济与金融研究院整理

5.5、基于周度 E-NELS 因子的主动量化策略研究

与月度主动量化研究框架一致，基于周度 E-NELS 因子，我们构建了两个主动量化选股模型，一个持仓 100 只股票，一个持仓 50 只股票，策略细节参见图表 76。

图表 76、基于周度 E-NELS 因子构建的主动量化策略框架

	策略一	策略二
差异性要求	1、每期选择 100 只股票； 2、调仓日：若当期持仓的股票下一期没有跌出前 500，仍然继续持有	1、每期选择 50 只股票； 2、调仓日：若当期持有的股票下一期没有跌出前 300，仍然继续持有
统一要求	选股池需删除 ST、同时保证上市天数满 180 天；交易成本：0.3%	

资料来源：兴业证券经济与金融研究院整理

从测试结果来看，基于周度 E-NELS 因子构建的主动量化策略表现同样优异，以选择 100 只股票的策略一为例：该策略年化超额收益率 29.3%，风险收益比高达 4.7。即便分年度来看，除 2017 年外，其他年份多头策略均能稳定的战胜基准。

图 77、针对于中证 500 的周度主动量化策略表现汇总

	年化收益率	波动率	风险收益比	胜率	回撤
中证 500 表现	1.0%	26.8%	0.04	54.2%	65.2%
策略一多头	30.5%	27.2%	1.12	57.7%	42.5%
策略二多头	32.9%	27.7%	1.19	58.5%	41.4%
策略一超额	29.3%	6.2%	4.71	61.7%	9.8%
策略二超额	31.8%	7.3%	4.38	60.4%	11.0%

资料来源：Wind，兴业证券经济与金融研究院整理

图 78、针对于中证 500 构建的周度主动量化策略分年度表现

	中证 500	策略一多头	策略二多头	策略一超额	策略二超额
2010 年	10.1%	50.2%	58.9%	36.2%	44.4%
2011 年	-33.8%	-20.5%	-19.1%	19.7%	21.9%
2012 年	0.3%	27.1%	28.8%	26.9%	28.8%
2013 年	16.9%	69.3%	76.1%	44.4%	50.4%
2014 年	39.0%	78.6%	71.4%	28.7%	23.8%
2015 年	43.1%	148.9%	164.3%	75.6%	86.4%
2016 年	-17.8%	9.4%	8.4%	32.5%	31.4%
2017 年	-0.2%	-8.6%	-6.7%	-8.3%	-6.4%
2018 年	-33.3%	-18.5%	-16.0%	22.5%	26.3%
2019 年	17.8%	31.8%	30.4%	12.0%	10.7%

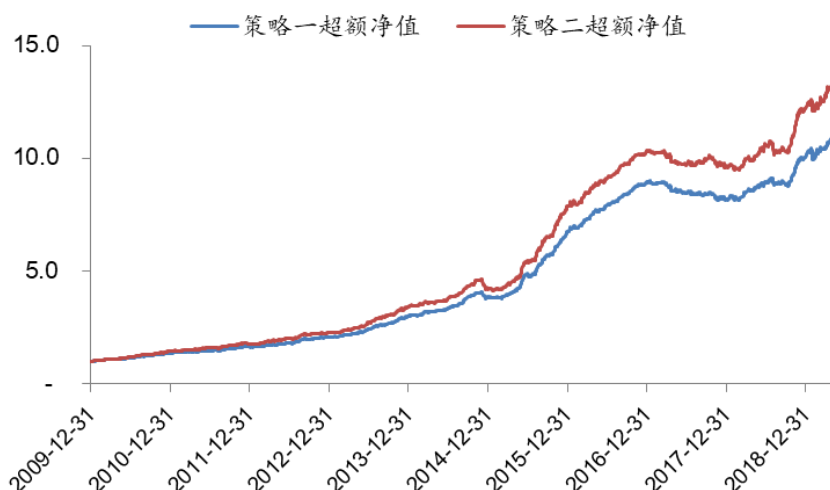
资料来源：Wind，兴业证券经济与金融研究院整理

图 79、针对于中证 500 构建的周度主动量化策略多头净值曲线



资料来源：Wind，兴业证券经济与金融研究院整理

图表 80、针对于中证 500 构建的周度主动量化策略超额净值曲线



资料来源：Wind，兴业证券经济与金融研究院整理

5.6、模拟组合表现汇总

从被动跟踪到指数增强，进一步到主动量化，不同策略的背后，一方面反映了对风险的逐步释放，另一面也体现了灵活性与特异性的不断提升。我们共计构建了 11 个不同类型的选股模型，全方位多角度的研究了周度和月度 E-NELS 因子在策略方面的可用性。每个策略的表现和应用场景不同，大家可以结合自己的需求选择使用。

图表 81、基于月度 E-NELS 因子构建的策略表现回顾

	年化收益	波动	夏普率	胜率	回撤
沪深 300 表现	0.2%	23.1%	0.01	51.1%	46.7%
全市场选择，跟踪误差 2%多头	8.0%	22.7%	0.35	52.4%	41.1%
全市场选择，跟踪误差 2%超额	7.7%	2.3%	3.34	59.8%	2.5%
全市场选择，跟踪误差 5%多头	8.3%	22.8%	0.37	52.6%	40.4%
全市场选择，跟踪误差 5%超额	8.0%	2.5%	3.17	59.0%	2.7%
全市场选择，指数成分股权重占比>80%多头	8.5%	22.8%	0.37	52.5%	39.8%
全市场选择，指数成分股权重占比>80%超额	8.2%	2.5%	3.31	59.8%	2.5%
中证 500 表现	1.0%	26.8%	0.04	54.2%	65.2%
全市场选择，跟踪误差 2%多头	13.6%	26.8%	0.51	55.2%	50.6%
全市场选择，跟踪误差 2%超额	12.6%	3.0%	4.27	62.0%	2.1%
全市场选择，跟踪误差 5%多头	17.4%	26.3%	0.66	55.7%	46.2%
全市场选择，跟踪误差 5%超额	16.2%	4.0%	4.08	59.9%	3.1%
全市场选择，指数成分股权重占比>80%多头	14.0%	26.5%	0.53	55.5%	46.8%
全市场选择，指数成分股权重占比>80%超额	12.7%	3.5%	3.59	59.6%	2.7%
主动量化，选择 100 只股票多头	26.1%	27.7%	0.94	56.7%	46.3%
主动量化，选择 100 只股票超额	25.3%	5.7%	4.44	61.2%	11.0%
主动量化，选择 50 只股票多头	27.0%	27.9%	0.97	57.1%	47.2%
主动量化，选择 50 只股票超额	26.1%	6.6%	3.93	60.2%	10.4%

资料来源：Wind，兴业证券经济与金融研究院整理

图表 82、基于周度 E-NELS 因子构建的策略表现回顾

	年化收益	波动	夏普率	胜率	回撤
中证 500 表现	1.0%	26.8%	0.04	54.2%	65.2%
增强模型多头	23.2%	25.3%	0.92	56.4%	42.1%
增强模型超额	21.6%	5.1%	4.26	60.7%	6.0%
主动量化选择 100 只股票多头	30.5%	27.2%	1.12	57.7%	42.5%

请务必阅读正文之后的信息披露和重要声明

主动量化选择 100 只股票超额	29.3%	6.2%	4.71	61.7%	9.8%
主动量化选择 50 只股票多头	32.9%	27.7%	1.19	58.5%	41.4%
主动量化选择 50 只股票超额	31.8%	7.3%	4.38	60.4%	11.0%

资料来源：Wind，兴业证券经济与金融研究院整理

6、总结

量化投资管理的未来会将走向何方？这是我们开篇提到的问题。人工智能第三次浪潮能否持续下去？这是大家普遍关注的问题。在本篇报告里，我们将量化投资与机器学习有机结合，给出了量化投资的新视角。

整个机器学习领域确实宛若汪洋，每种算法都包罗万象。我们这次切入时，并没有从广度入手，将所有的算法都做一个尝试，而是以 Adaboost 算法为切入点，深入探讨了传统 Adaboost 与改进 Adaboost 之间的差异。进一步地，以改进版 Adaboost 算法为抓手，构建了月度、周度选股模型。当然文章也重点探讨了行业市值影响、风格切换分析、并结合不同的场景进行相应的研究。最终在周度和月度因子的基础上，构建了多种选股策略，供大家参考。

人工智能在量化投资领域的应用之路也许不会一帆风顺，但我们的努力绝不会停止，我们将坚守卖方研究的初衷，砥砺前行，争取为大家提供更多有价值的成果。

7、附录

图表 83、部分因子释义

因子	释义	分类	IC	T 值
DP_LTM	过去一年现金分红 / 总市值	价值	0.02	2.4
EP_TTM	净利润(不含少数股东损益)_TTM / 总市值	价值	0.03	2.4
EP_FY1	一致预期净利润_FY1 / 总市值	价值	0.05	4.1
EP_Fwd12M	每股收益_未来 12 个月预测值 / 收盘价	价值	0.05	4.1
EP_TTM_Deducted	扣非经常性损益后净利润_TTM / 总市值	价值	0.02	2.2
OCFP_TTM	经营活动现金流量净额_TTM / 总市值	价值	0.02	3.7
SP_TTM	营业收入_TTM / 总市值	价值	0.03	3.4
SP_LYR	营业收入_最新年报 / 总市值	价值	0.03	3.0
Revenue_LYR_YoY	营业收入_最新年报的同比增长率	成长	0.01	-0.8
NetProfit_LYR_YoY	净利润_最新年报的同比增长率	成长	0.00	-0.6
Revenue_SQ_YoY	单季度营业收入同比增长率	成长	0.03	4.1
Revenue_SQ_QoQ	单季度营业收入环比增长率	成长	0.01	2.4
Revenue_TTM_YoY	营业收入_TTM/去年营业收入_TTM-1	成长	0.01	1.5
NetProfit_SQ_YoY	单季度净利润同比增长率	成长	0.04	5.2
NetProfit_SQ_QoQ	单季度净利润环比增长率	成长	0.01	2.4
NetProfit_TTM_YoY	净利润_TTM/去年净利润_TTM-1	成长	0.01	1.5
Asset_LR_YoY	总资产同比变动	成长	0.01	-0.7
Equity_LR_YoY	净资产(股东权益)同比变动	成长	0.01	-1.0
RTN_1D	复权收盘价 / 复权收盘价_1 天前 - 1	动量反转	0.02	1.7
RTN_5D	复权收盘价 / 复权收盘价_5 天前 - 1	动量反转	0.05	4.7
RTN_20D	复权收盘价 / 复权收盘价_20 天前 - 1	动量反转	0.07	6.0
EPS_FY0_R1M	一致预期 EPS_FY0 过去 20 天的变化率	情绪	0.03	7.4
EPS_FY0_R3M	一致预期 EPS_FY0 过去 60 天的变化率	情绪	0.03	4.6
ROE_FY0_R1M	一致预期 ROE_FY0 过去 20 天的变化率	情绪	0.03	7.8
ROE_FY0_R3M	一致预期 ROE_FY0 过去 60 天的变化率	情绪	0.02	4.4
Rating_R1M	分析师综合评级 1 个月的变化率	情绪	0.02	5.3
Rating_R3M	分析师综合评级 3 个月的变化率	情绪	0.02	5.2
Rating	分析师综合评级	情绪	0.02	2.8
TargetReturn	一致预测目标价 / 收盘价 - 1	情绪	0.05	4.6
ROE_FY0	一致预期净利润_ / 股东权益合计_最新财报	质量	0.01	0.06
CurrentRatio_LR	流动资产_最新财报 / 流动负债合计_最新财报	质量	0.01	0.05
LTDebt2Equity_LR	非流动负债_最新财报 / 股东权益合计_最新财报	质量	0.01	0.10
IVR_20D	1 - Fama-French 回归的调整 r 方	另类	0.08	11.5
IVFF_20D	Fama-French 回归残差项的年化波动率	另类	0.09	8.9
Close_Ln	收盘价的自然对数	另类	0.04	2.7
MaxReturn_20D	过去 20 天日收益率的最大值	另类	0.07	7.0
Amount_20D_Avg	过去 20 天日均成交额	另类	0.11	8.6
Amount_240D_Avg	过去 240 天日均成交额	另类	0.07	5.8
Turnover_20D_Avg	过去 20 天日均换手率	另类	0.08	6.0
Turnover_240D_Avg	过去 240 天日均换手率	另类	0.03	2.7
SmallTradeFlow_1D	散户量差(仅主动) / 成交量	另类	0.04	5.6
BuyMinusSell_5D_V	主买总量 - 主卖总量 过去 5 天的和	另类	0.01	1.1
Num_EPS_FY0	EPS 预测值的数量	另类	0.00	-0.2

资料来源：Wind，兴业证券经济与金融研究院整理

风险提示：本报告模型及结论全部基于对历史数据的分析，当市场环境变化时，存在模型失效风险。

分析师声明

本人具有中国证券业协会授予的证券投资咨询执业资格并注册为证券分析师，以勤勉的职业态度，独立、客观地出具本报告。本报告清晰准确地反映了本人的研究观点。本人不曾因，不因，也将不会因本报告中的具体推荐意见或观点而直接或间接收到任何形式的补偿。

投资评级说明

投资建议的评级标准	类别	评级	说明
报告中投资建议所涉及的评级分为股票评级和行业评级(另有说明的除外)。评级标准为报告发布日后的12个月内公司股价(或行业指数)相对同期相关证券市场代表性指数的涨跌幅。其中：A股市场以上证综指或深圳成指为基准，香港市场以恒生指数为基准；美国市场以标普500或纳斯达克综合指数为基准。	股票评级	买入	相对同期相关证券市场代表性指数涨幅大于15%
		审慎增持	相对同期相关证券市场代表性指数涨幅在5%~15%之间
		中性	相对同期相关证券市场代表性指数涨幅在-5%~5%之间
		减持	相对同期相关证券市场代表性指数涨幅小于-5%
		无评级	由于我们无法获取必要的资料，或者公司面临无法预见结果的重大不确定性事件，或者其他原因，致使我们无法给出明确的投资评级
	行业评级	推荐	相对表现优于同期相关证券市场代表性指数
		中性	相对表现与同期相关证券市场代表性指数持平
		回避	相对表现弱于同期相关证券市场代表性指数

信息披露

本公司在知晓的范围内履行信息披露义务。客户可登录 www.xyzq.com.cn 内幕交易防控栏内查询静默期安排和关联公司持股情况。

使用本研究报告的风险提示及法律声明

兴业证券股份有限公司经中国证券监督管理委员会批准，已具备证券投资咨询业务资格。

本报告仅供兴业证券股份有限公司(以下简称“本公司”)的客户使用，本公司不会因接收人收到本报告而视其为客户。本报告中的信息、意见等均仅供客户参考，不构成所述证券买卖的出价或征价邀请或要约。该等信息、意见并未考虑到获取本报告人员的具体投资目的、财务状况以及特定需求，在任何时候均不构成对任何人的个人推荐。客户应当对本报告中的信息和意见进行独立评估，并应同时考量各自的投资目的、财务状况和特定需求，必要时就法律、商业、财务、税收等方面咨询专家的意见。对依据或者使用本报告所造成的一切后果，本公司及/或其关联人员均不承担任何法律责任。

本报告所载资料的来源被认为是可靠的，但本公司不保证其准确性或完整性，也不保证所包含的信息和建议不会发生任何变更。本公司并不对使用本报告所包含的材料产生的任何直接或间接损失或与此相关的其他任何损失承担任何责任。

本报告所载的资料、意见及推测仅反映本公司于发布本报告当日的判断，本报告所指的证券或投资标的的价格、价值及投资收入可升可跌，过往表现不应作为日后的表现依据；在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的报告；本公司不保证本报告所含信息保持在最新状态。同时，本公司对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

除非另行说明，本报告中所引用的关于业绩的数据代表过往表现。过往的业绩表现亦不应作为日后回报的预示。我们不承诺也不保证，任何所预示的回报会得以实现。分析中所做的回报预测可能是基于相应的假设。任何假设的变化可能会显著地影响所预测的回报。

本公司的销售人员、交易人员以及其他专业人士可能会依据不同假设和标准、采用不同的分析方法而口头或书面发表与本报告意见及建议不一致的市场评论和/或交易观点。本公司没有将此意见及建议向报告所有接收者进行更新的义务。本公司的资产管理部门、自营部门以及其他投资业务部门可能独立做出与本报告中的意见或建议不一致的投资决策。

本报告并非针对或意图发送予或为任何就发送、发布、可得到或使用此报告而使兴业证券股份有限公司及其关联子公司等违反当地的法律或法规或可致使兴业证券股份有限公司受制于相关法律或法规的任何地区、国家或其他管辖区域的公民或居民，包括但不限于美国及美国公民(1934年美国《证券交易所》第15a-6条例定义为本「主要美国机构投资者」除外)。

本报告的版权归本公司所有。本公司对本报告保留一切权利。除非另有书面显示，否则本报告中的所有材料的版权均属本公司。未经本公司事先书面授权，本报告的任何部分均不得以任何方式制作任何形式的拷贝、复印件或复制品，或再次分发给任何其他人，或以任何侵犯本公司版权的其他方式使用。未经授权的转载，本公司不承担任何转载责任。

特别声明

在法律许可的情况下，兴业证券股份有限公司可能会持有本报告中提及公司所发行的证券头寸并进行交易，也可能为这些公司提供或争取提供投资银行业务服务。因此，投资者应当考虑到兴业证券股份有限公司及/或其相关人员可能存在影响本报告观点客观性的潜在利益冲突。投资者请勿将本报告视为投资或其他决定的唯一信赖依据。

兴业证券研究

上 海	北 京	深 圳
地址：上海浦东新区长柳路36号兴业证券大厦15层	地址：北京西城区锦什坊街35号北楼601-605	地址：深圳市福田区皇岗路5001号深业上城T2座52楼
邮编：200135	邮编：100033	邮编：518035
邮箱：research@xyzq.com.cn	邮箱：research@xyzq.com.cn	邮箱：research@xyzq.com.cn